$2.45 ACADEMIC PAPERBACKS

Geometric Transformations

Volume 2: Projective Transformations

P. S. Modenov

A. S. Parkhomenko

Geometric Transformations

P. S. MODENOV and A. S. PARKHOMENKO

VOLUME 2

Projective Transformations

Translated and adapted from the first Russian edition by

MICHAEL B. P. SLATER

Published in cooperation with the
SURVEY OF
RECENT EAST EUROPEAN MATHEMATICAL LITERATURE
A project conducted by
ALFRED L. PUTNAM AND IZAAK WIRSZUP
Department of Mathematics,
The University of Chicago, under a
grant from the National Science Foundation

ACADEMIC PRESS New York and London

First published in the Russian language under the title
Geometricheskie Preobrazovaniya
in 1961 by Izdatel'stvo Moskovskogo Universitet,
Moscow, U.S.S.R.

ACADEMIC PRESS INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS INC. (LONDON) LTD.
Berkeley Square House, London W.1

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 65-25004

PRINTED IN THE UNITED STATES OF AMERICA

Preface to Volume 2 of the English Edition

This is the second volume of a two-volume translation of the Russian book *Geometric Transformations,* by Modenov and Parkhomenko. It contains a translation of Chapters V (with Appendix) and VI of the original, here Chapters I and II, respectively.

The greater portion of this volume is concerned with projective transformations. These are the collinearity-preserving transformations of the projective plane. It turns out that many of the mappings of the ordinary plane that preserve collinearity may best be regarded as defined on an extended object, called the projective plane. In particular, the affine transformations that were considered in Chapter IV of the first volume *(Euclidean and Affine Transformations)* may very conveniently be regarded as those projective transformations which fix the *ideal line.*

Chapter I therefore starts with the motivation of the construction of the projective plane, followed by a number of alternative constructions for it. Later most of the basic facts are proved, and some of the applications outlined.

Chapter II deals with an independent topic, but at the same level of sophistication. However, the appendices to Chapter I refer to more advanced concepts, and these are not motivated or treated in detail. The unprepared reader should not be afraid to read through them, picking up what he can, and taking the rest on faith. They serve as distant glimpses of some of the methods and concerns of modern geometry and algebra, respectively.

The concrete details of proofs and discussions (except in the appendices) are mostly very elementary, and often take place within the Euclidean plane or Euclidean space. The reader might profit also from a smattering of knowledge of the theory of linear equations, including familiarity with determinants. However, he does need a little of the more intangible prerequisite known as *maturity*, including the ability to recognize a proof when he sees one, even though some of the concepts involved

may be unfamiliar. It would be advisable also for him to have some acquaintance with the contents of Volume 1, at least to the point of being at home with the concept of a group of transformations of a given set. While comparatively few specific results from Volume 1 are actually needed, this second volume is, of course, in the same spirit, and an unprepared undergraduate who plunges straight into it will probably feel a little uncomfortable.

In general, this book is probably at about the right level for an undergraduate reader. However, interest and ability are more relevant for profiting from it than is quantity of previous knowledge.

The Russian authors suggest that their book can best serve as extracurricular material for geometry seminars in universities and teacher-training colleges, as extra background material for high school teachers, and as source material for school mathematics clubs (under a teacher's guidance). Apart from the last category, for which this volume may be a little advanced, the same might be suggested for the American translation.

Chicago, Illinois M. SLATER
1965

Translator's Note

The translation is quite free. Although it retains all of the original text, it recasts many passages and in several sections includes additional background discussion and motivation. Apart from Appendix 2 to Chapter I, however, the section headings are the same, and in the same order, as in the original. Wherever I have added to or changed the text, I have tried to remain consistently within its spirit. The burden of responsibility for all deviations from the original must, of course, rest entirely on me.

M. S.

This page intentionally left blank

Preface to the Russian Edition

This book is intended for use in geometry seminars in universities and teacher-training colleges. It may also be used as supplementary reading by high school teachers who wish to extend their range of knowledge. Finally, many sections may be used as source material for school mathematics clubs under the guidance of a teacher.

The subject matter is those transformations of the plane that preserve the fundamental figures of geometry: straight lines and circles. In particular, we discuss orthogonal, affine, projective, and similarity transformations, and inversions.

The treatment is elementary, though in a number of instances (where a synthetic treatment seems more cumbersome) coordinate methods are used. A little use is also made of vector algebra, but the text here is self-contained.

In order to clarify a number of points, we give some elementary facts from projective geometry; also, in the addendum to Chapter I of Volume 2 (the topology of the projective plane), the structure of the projective plane is examined in greater detail.

The authors feel obliged to express their thanks to Professor V. G. Boltyanskii, who carefully read the manuscript and made a number of valuable suggestions. They also wish to thank Miss V. S. Kapustina for editing the manuscript and removing many inadequacies of presentation. They would like also to say that Chapter II of Volume 2 was written with the help of an article on inversion written by V. V. Kucherenko, a second-year physics student. It is to him that we owe the elegant proof of the fundamental theorem that any circle transformation can be represented as the product of an inversion and a similarity transformation, and also as the product of an inversion and a rotation (or a reflection).

Moscow THE AUTHORS
January 1961

This page intentionally left blank

Contents

This page intentionally left blank

CHAPTER I

Projective Transformations

I. The Concept of a Projective Plane

We have defined an affine mapping of one plane onto another as a one-one mapping that preserves collinearity. The question arises as to whether there are correspondences between planes that preserve collinearity and yet are not affine. Without further restriction, we can get some very uninteresting mappings; for example, the mapping of π into π' that takes every point of π into some point of a given line l of π'. It is clear that nothing useful can be said about such a mapping. In this chapter, we consider some very important mappings of one plane π into another that preserve collinearity and are also one-one but that are not defined in all π and do not have every point of π' in their range.

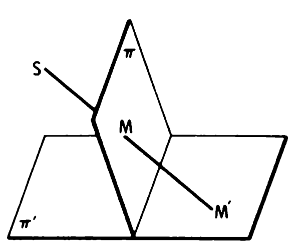

Fig. 1

Consider first the example of a *perspectivity*. Let π and π' be two distinct planes (Fig. 1), and let S be any point not lying on either of them. We project the plane π into π' through the point S. That is to say, given a point M of π, we make correspond to it the point M' (if there is one) in which

SM meets π'. We call this mapping a *perspectivity* between π and π', and S the *center of perspective.*

If n is any line of π, its image under the perspectivity is the line in which π' intersects the plane through S and n. Thus perspectivities preserve collinearity. However, perspectivities are not, in general, defined on all of π, nor is the image of π under the perspectivity all of π'. For the line k of intersection of π with the plane through S parallel to π' has no image (that is, the mapping is not defined on any point of k), and no point of the line l' of intersection of π' with the plane through S parallel to π is the image of any point of π under the perspectivity (Fig. 2). We may define a perspectivity as a certain kind of line-preserving map of one plane (with a line deleted) onto another plane (with a line deleted). Note that there are no lines k and l', if and only if π and π' are parallel. In this case, every perspectivity between π and π' is an affine mapping

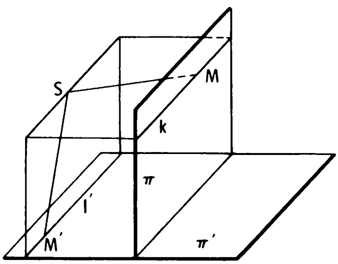

Fig. 2

As a familiar example of a perspectivity, consider a photograph of the land (considered as a horizontal plane). This is the perspectivity between the horizontal plane of the land and the vertical plane (or part of it) of the photographic film. The center of perspective is the lens of the camera. Imagine, for example, a photograph of level ground with a railway running along it. The image of the figure of the two rails, which have no point in common, is a pair of lines which do have a point in common. The point of intersection of the lines in the photograph is a point of the image having no inverse image in the original.

It seems natural to supply the planes π and π' with new points that will make the perspectivity defined on all of π and make it have all of π' as range. To see how this should be done, consider any point T of k, and let n be a line through

it other than k. Consider the images under the perspectivity of a sequence of points on n approaching ever nearer to T. Their images in π' will be a sequence of collinear points tending to infinity. Similarly, a sequence of points tending to T along n from the other side will be mapped into points of the same line n' of π', this time tending to infinity in the other direction. It seems natural, then, to supply π' with a new "point" T' lying on n' and infinitely far away on n' but not thought of as lying on one end rather than the other. Then the image of the ordinary line n will be a new object n' together with T', which we will call a line. We thus have to assign a new point "at infinity" to each line of π', corresponding to the point in which the inverse image of this line in π cuts k. Similarly (since the inverse of a perspectivity is clearly a perspectivity), we must assign to each line of π a new "point" lying "at infinity" on this line and regard this as the "point" mapped by the perspectivity into the point of intersection of the image line with l'.

Suppose now that n_1 and n_2 are two parallel lines of π. Their images will be two lines of π' that meet on l' (or are both parallel to it). Thus we need to assign the *same* "point at infinity" to each of two parallel lines. Similarly, the images of two lines of π that intersect on k are two parallel lines of π'. Thus (to preserve the property of collinearity) we want this point of intersection to lie on both of the image lines; that is, they must have the same "point at infinity." Finally, the image of the line k is the set of all the points at infinity of π', and since we want to preserve collinearity, we will have to say that all the points at infinity of π' are collinear on the "line at infinity."

So far, everything we have said is heuristic, and the reader may get a sense of unreality from all these "points at infinity," which do not "really" exist. However, the preceding discussion will clarify what we are driving at in the next part of this section, where there will be no quotation marks. What we shall do is to construct a new object out of the plane by adding what we have called here the "points at infinity," and in the

new, enriched object, which we call the projective plane, *define* what we mean by a line, the point of intersection of two lines, and so on.

1.1. First Model of the (Real) Projective Plane

We call the set of all lines of the plane parallel to a given line a *sheaf* of parallel lines. Thus there is one sheaf for each angle α $(0 \leq \alpha < \pi)$. With each sheaf we associate a new element, whose nature is of no interest to us, but in such a way that different sheafs have different objects associated with them. As an example, we could associate with each sheaf the angle it makes with the x axis in some coordinate system. Each of these new objects may be called a "point at infinity" or an "improper" point. We will call them *ideal* points. The set of all points, ordinary and ideal, we call the (real) *projective plane.*

We next define a *projective line.* We say that a subset of the projective plane is a *line* (or, if we wish to avoid confusion with ordinary lines, a *projective line*) if it consists of all the points of an ordinary line, together with the ideal point of the sheaf to which that ordinary line belongs, and no other points. In addition, the set of all the ideal points (and no ordinary points) is also defined to be a projective line. It may be called the "line at infinity," or the "improper" line, but we will call it the *ideal* line.

Note that our definition makes sense, since a given ordinary line belongs to one and only one sheaf—the sheaf of lines parallel to it.

We now define two (projective) lines to *intersect* in the point P, provided that P is a point common to both of them (remember lines were defined simply as being certain sets of points). Similarly, we say that the point P *lies on* l, or that l *passes through* P, if and only if P is a member of l.

We show now that the projective plane satisfies a fundamental theorem of the ordinary plane, and also its dual (obtained by interchanging "line" with "point" and "lies on" with "passes through"):

Theorem 1. *Through any two distinct points of the projective plane there passes one and only one (projective) line.*

Proof of 1. Suppose first that both the points are ordinary. Then there is a unique ordinary line passing through them both, and the unique projective line of which this ordinary line is part also passes through both of them. We leave the proof that this is the only projective line passing through these two points to the reader.

Suppose next that one of the points A is ordinary and that the other B is ideal. There is a unique (ordinary) line of the sheaf corresponding to B that passes through A—the line through A and in the appropriate direction. This ordinary line, together with the point B, is a projective line through A and B. Once again we leave the proof of uniqueness to the reader.

Suppose finally, that both points are ideal. Then the ideal line passes through them, and since it is the only projective line containing more than one ideal point, it is the only projective line passing through the two points. ▼

Theorem 2. *Any two distinct lines have one and only one point of intersection.*

This is the dual statement to 1, and is not true in the ordinary plane. It is this fact that makes projective geometry (the study of the projective plane) a more elegant subject than affine geometry (the study of the ordinary, or affine, plane). In the plane, two lines cannot intersect in more than one point, but it is possible for two lines not to intersect at all, and such exceptional pairs of lines (parallel lines) play an important part in Euclidean geometry.

Proof of 2. If both the lines are not ideal, then the ordinary lines that form part of them either intersect or are parallel. If they intersect, the point of intersection is a point of intersection of the projective lines too. If they are parallel, then the projective lines have a common ideal point, since then both the ordinary lines belong to the same sheaf.

If one of the lines is the ideal line, then it has in common with the other the ideal point on the latter.

That no two distinct lines intersect in more than one point follows at once from the uniqueness part of Theorem 1. ▼

We see that there is no such concept as parallelism; any two lines intersect.

Let us come back now to our perspectivity. We will see that it can be tidied up if we define a perspectivity for two *projective* planes. Let π and π' be two affine planes, let Π and Π' be the projective planes obtained from them by adding ideal points, and let α be the perspectivity between π and π' with center S. We define the mapping A of Π onto Π' to be α on the points of π other than the line k. Given any ideal point P of Π, let n be any projective line through it and let the image n' of n meet l' in P' (here k and l' are as in Fig. 2). Then we set $A(P) = P'$. If Q is any point of k, let n be any line of π through it and n' its image under α. Then we set $A(Q) = Q'$, where Q' is the ideal point on the projective line of Π' of which n' is a part. The reader should verify that these are both good definitions in the sense that they give the same point P' or Q', whatever our choice of n. He should verify further that A is a one-one mapping of Π onto Π', and that it preserves collinearity, that is, that the image of a projective line under A is a projective line. In Section 2 of this volume we shall give another way of defining a perspectivity between two projective planes. At any rate, it is clear that it is more natural to define a perspectivity as a mapping from one projective plane onto another, rather than from part of one affine plane onto part of another.

1.2. Homogeneous Coordinates

The method of coordinates in the Euclidean plane can be used with great success for the solution of a wide range of questions in Euclidean geometry. It can also be used on the

projective plane. However, the existence of ideal points complicates the process of introducing coordinates in the projective plane.

Let Π be the projective plane obtained in the way we described from the affine plane π. We introduce in π any system of Cartesian coordinates (not necessarily rectangular). If M is an ordinary point of Π, then in this system of coordinates it will have a well-defined pair of coordinates x, y.

Consider the ordered triple of numbers $(x, y, 1)$, and the class of all triples (x_1, x_2, x_3) proportional to it:

$$x_1 : x_2 : x_3 = x : y : 1,$$

that is,

$$x_1 = \lambda x, \qquad x_2 = \lambda y, \qquad x_3 = \lambda,$$

where λ takes any real value except zero. Since $x_3 \neq 0$,

$$x = x_1/x_3, \qquad y = y_2/y_3.$$

Thus each triple of real numbers, the third of which is not zero, corresponds to a unique point M of π. If $M = M(x, y)$, then any triple (x_1, x_2, x_3) proportional to $(x, y, 1)$ (in the sense just explained) is called a triple of *homogeneous coordinates for M* (in the given coordinates system), or we say simply that *M has homogeneous coordinates* (x_1, x_2, x_3). Note, however, that the homogeneous coordinates of M are not unique; any triple (x_1', x_2', x_3') proportional to (x_1, x_2, x_3) serves just as well.

Suppose now that M is an ideal point of Π. Then M is associated with the set of all those lines of π that are parallel to some given vector $\mathbf{a}$. Suppose that $\mathbf{a}$ has coordinates (X, Y) (in the given coordinate system), and consider the collection of all triples (x_1, x_2, x_3) proportional to the triple $(X, Y, 0)$:

$$x_1 : x_2 : x_3 = X : Y : 0,$$

that is,

$$x_1 = \lambda x, \qquad x_2 = \lambda y, \qquad x_3 = 0,$$

where λ can take any real value except zero.

We say any triple of this collection is a triple of *homogeneous coordinates for M*. If we had chosen a different vector **b** to which every line of the sheaf associated with M is parallel, then the coordinates (X', Y') of **b** would be proportional to those of **a**, so that the class of triples proportional to $(X', Y', 0)$ would be the same as the class of triples proportional to $(X, Y, 0)$. So our definition for the collection of all homogeneous coordinates of M is a good one. It is easy to see also that different ideal points M, M' have different classes of homogeneous coordinates, and that every class of homogeneous coordinates of which the third is zero is associated with some ideal point. Indeed, the triple $(X, Y, 0)$ is associated with the sheaf of lines parallel to the vector whose coordinates are (X, Y), and with no other, and thus with one and only one special point.

We have thus shown that, after proportional triples have been identified, there is one and only one triple corresponding to each point of Π, and, conversely, there is one and only one point of Π corresponding to each triple, except for the triple $(0, 0, 0)$.

One of the fundamental facts of analytic geometry is that the equation of any line of the plane is linear, that is, of the first degree in the coordinates, and, conversely, that the set of points of the plane that satisfy any given equation linear in x and y is a line. We have an analogous theorem for the projective plane:

Theorem 3. *Any (projective) line of the projective plane has a homogeneous linear equation corresponding to it. Conversely, any linear homogeneous equation in three unknowns determines a projective line.*

This theorem means that given a projective line l, there are real constants A, B, C, not all zero, such that l consists precisely of those points $M(x_1, x_2, x_3)$ whose coordinates x_1, x_2, x_3 satisfy the equation

$$Ax_1 + Bx_2 + Cx_3 = 0. \tag{1}$$

Note that if the triple (x_1, x_2, x_3) satisfies Eq. (1), so does any proportional triple, so that it does not matter what triple of homogeneous coordinates we take for M.

The second part of the theorem states that, conversely, the set of those points of which the homogeneous coordinates (that is, all possible triples of homogeneous coordinates) satisfy an equation of the form (1)—with not all of A, B, C zero—is a projective line.

Proof. Let L be any projective line. If L is the ideal line, it has the equation

$$x_3 = 0.$$

For any ideal point has a third coordinate 0, and, conversely, any point with the third coordinate 0 is ideal. Note that if the third coordinate is 0 in one triple of homogeneous coordinates for a point M, then it is 0 in every such triple.

If L is not the ideal line, then the ordinary equation (in the given coordinate system) of the set of all the ordinary points on it is of the form

$$Ax + By + C = 0, \tag{2}$$

where either $A \neq 0$ or $B \neq 0$. For, by the definition of a projective line (other than the special one), the set of ordinary points on it is an ordinary line.

If M is any ordinary point of L, and (x, y) are its ordinary coordinates, then any triple (x_1, x_2, x_3) of homogeneous coordinates for M is such that

$$x_3 \neq 0, \qquad x = x_1/x_3, \qquad y = x_2/x_3,$$

and, therefore,

$$Ax_1/x_3 + Bx_2/x_3 + C = 0,$$

so that

$$Ax_1 + Bx_2 + Cx_3 = 0.$$

Thus *any* triple of homogeneous coordinates for M satisfies Eq. (1), and, conversely, no ordinary point not on L can have

homogeneous coordinates satisfying it. To prove this, write out the above argument backwards; it will show that any ordinary point whose coordinates satisfy (1) must lie on the line (2). Suppose next that M is the ideal point of L. Then the vector $\mathbf{a} = (-B, A)$ is parallel to the ordinary line l which is part of L, so that any triple of homogeneous coordinates (x_1, x_2, x_3) for M are such that

$$x_1 : x_2 : x_3 = -B : A : 0,$$

or

$$x_3 = 0, \qquad x_1 A = x_2(-B).$$

But then

$$Ax_1 + Bx_2 + Cx_3 = 0,$$

so that Eq. (1) is satisfied by any triple of homogeneous coordinates for M. Thus any triple of homogeneous coordinates for *any* point M lying on L satisfies Eq. (1).

Conversely, any point M for which a triple (and hence every triple) of homogeneous coordinates satisfies (1) must lie on L. We have already indicated a proof for ordinary points, and if M is ideal, then $x_3 = 0$, so that the coordinates (x_1, x_2, x_3) of M satisfy $Ax_1 + Bx_2 = 0$, that is, are proportional to $(-B, A, 0)$. So M is the ideal point of L; that is, it lies on L.

We now prove the converse, that any equation (1) in which not all of A, B, C are zero is an equation for a projective line. If A and B are both zero, then $C \neq 0$, and the equation reduces to $x_3 = 0$, the equation of the ideal line. Suppose now that one of A and B is not zero, and consider the set of ordinary points satisfying (1). Since for ordinary points $x_3 \neq 0$, these points will also satisfy

$$Ax_1/x_3 + Bx_2/x_3 + C = 0,$$

or

$$Ax + By + C = 0,$$

where $x = x_1/x_3$, $y = x_2/x_3$ are the ordinary coordinates of these points. Thus, the set of all ordinary points of Π satisfying (1) is an ordinary line l.

We now find the ideal points satisfying (1). Since we must have $x_3 = 0$, $Ax_1 + Bx_2 = 0$, so that an ideal point satisfying (1) must be the point $(-B, A, 0)$. But this is just the ideal point associated with l. So we have proved that (1) represents the projective line obtained by adding an ideal point to l. ▼

1.3. SECOND MODEL OF THE PROJECTIVE PLANE

Let π be an ordinary plane, and let Π be the projective plane obtained from it in the usual way. Let S be any point of space not on π. We make correspond to each ordinary point M of Π the ray SM. If M is an ideal point of Π, it is associated with a sheaf of parallel lines of π, all parallel to a given vector $\mathbf{a}$. We associate with M the ray through S parallel to $\mathbf{a}$.

We have thus constructed a correspondence between the points of a projective plane Π and the rays through a point S of space. The correspondence is one-one and onto; to each point of Π there corresponds one and only one ray of S, and to each ray through S there corresponds one and only one point of Π.

What sort of a geometric object in space corresponds under this mapping to a line L of Π? Suppose first that L is not the ideal line. Consider the plane L_0 through S and the ordinary points of L. (Since these ordinary points comprise an ordinary line, there is such a plane.) Then each point of L corresponds to a ray through S lying in the plane L_0; in fact, it corresponds to the ray LM, if M is an ordinary point, and to the ray through S parallel to the ordinary part of L, if M is the ideal point of L. Conversely, any ray lying in the plane L_0 corresponds to a point of L.

Suppose next that L is the ideal line. Then the set of rays corresponding to the points of L comprise the plane L_0 through S and parallel to π.

Thus we have established a one-one correspondence between the rays through S in space and the points of Π, such that a

line of Π corresponds to a plane through S, and vice versa.

The notions of incidence carry over to this model unchanged. Thus, we speak of a ray (through S) *lying on* a plane through S, and this holds if and only if the corresponding point of Π lies on the corresponding line of Π. We talk of the plane (through S) *passing through* two given rays through S, and this too occurs if and only if the corresponding line of Π passes through the corresponding two points of Π.

These considerations give us ground for considering the set of all rays through S to *be* the projective plane, the "points" of the plane being the rays through S. With these new definitions of what we mean by the projective plane, its points, and its lines, the two fundamental incidence properties carry over:

1. *Through any two distinct "points" of the projective "plane" passes one and only one projective "line."*

2. *Any two distinct projective "lines" have one and only one projective "point" in common.*

These statements may easily be verified; for example, the second says that two planes through S intersect in a unique ray through S. However, they are also consequences of the isomorphism that we established between this model and the first model of the projective plane. All we had to do was translate these statements from the first model, and since they were true there, they must be true here.

1.4. Projective Coordinates

We now show how we may associate coordinates with the "points" of the second model.

Let $Sx_1x_2x_3$ be any Cartesian coordinate system (not necessarily rectangular) with origin S (Fig. 3). Given any "point", that is, any ray through S, choose any point M on the ray and suppose that its coordinates are (x_1, x_2, x_3). Then we say that the "point" *has homogeneous coordinates* (x_1, x_2, x_3).

If we had chosen another point M' of the ray, with (x_1', x_2', x_3') as coordinates, then

$$x_1':x_2':x_3' = x_1:x_2:x_3,$$

and, conversely, if

$$x_1':x_2':x_3' = x_1:x_2:x_3,$$

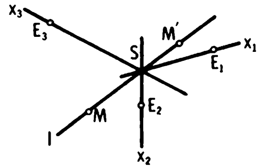

Fig. 3

then the point $M'(x_1', x_2', x_3')$ lies on the given ray. Thus any triple proportional to (x_1, x_2, x_3) is also a triple of homogeneous coordinates for the ray, and, conversely, no triple not proportional to (x_1, x_2, x_3) can be a triple of homogeneous coordinates for it.

Thus we have made correspond to each "point" of this model of the projective plane a complete class of proportional triples, any one of which we regard as being the homogeneous coordinates of that "point". Conversely, any class of proportional triples is associated with some "point"; for example, the class of triples proportional to a given triple (x_1, x_2, x_3) is associated with the "point" that is the ray SM, where M is the point of space with coordinates (x_1, x_2, x_3). Note that in all this we exclude S itself; the triple (0, 0, 0) does not figure in any of our classes, and we do not count S as being a typical point M of any of our rays.

Since the equation of any plane through the origin S has an equation of the form

$$Ax_1 + Bx_2 + Cx_3 = 0,$$

and since, conversely, the locus of points whose coordinates satisfy such an equation is a plane through S, we arrive at the following theorem, just as in the case of the first model of the projective plane

Theorem 4. *Any "line" of the projective plane (second model) has a homogeneous linear equation $Ax_1 + Bx_2 + Cx_3 = 0$, where one at least of A, B, C is nonzero. Conversely, the set of "points" whose coordinates satisfy such an equation is a "line."*

1.5. Connection between Projective Coordinates in the First and Second Models of the Projective Plane

Let us choose, as before, a system of coordinates $Sx_1x_2x_3$ in space with origin S. Let π be the plane through $E_3(0, 0, 1)$ parallel to the plane x_1Sx_2. We introduce coordinates in the plane π by taking E_3 for the origin and the lines E_3x and E_3y, respectively parallel to Sx_1 and Sx_2, for the coordinate axes (Fig. 4). We select the same scales on E_3x and E_3y as we have along Sx_1 and Sx_2 in the given system. Let Π be the projective plane obtained from π in the usual way. Then the projective coordinates of any "point" l are homogeneous coordinates of the point M of Π corresponding to l under the construction of Section 1.3 above.

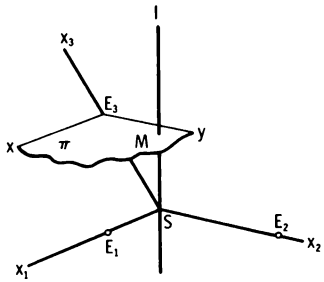

Fig. 4

To prove this, suppose first that M is an ordinary point and that its coordinates are $(x_1, x_2, 1)$. Then its coordinates in the coordinate system we introduced for π are (x_1, x_2). This follows from the way we chose our axes and unit points in π. Now l is the line SM, so that it has the projective coordinates $(x_1, x_2, 1)$, since these are the coordinates of a point M on it. On the other hand, the homogeneous coordinates of M, which has ordinary coordinates (x_1, x_2), are $(x_1, x_2, 1)$, or any multiple. We see that M and l have the same coordinates.

If M is ideal, then l is parallel to π and so lies in the plane $x_3 = 0$ (since π was constructed to be parallel to this plane). If P is any point on l, its coordinates $(x_1, x_2, 0)$, then (x_1, x_2) are the coordinates of the vector associated with M, but this means that the coordinates of M are also $(x_1, x_2, 0)$. We thus see that in this case too the projective coordinates of the "point" l are the same as the homogeneous coordinates of the point M of Π to which it corresponds.

1.6. Comments

We have given two models of the projective plane. We now briefly sketch two more.

Third Model. We take the projective plane to be the surface of the unit sphere, the "points" to be the pairs of antipodal points, and the "lines" to be the great circles. A "line" is a set of "points", for if one of two antipodal points lies on a great circle, so does the other. The two fundamental theorems hold: there is a unique "line" through any two distinct "points", and any two distinct "lines" intersect in a unique "point." The first of these propositions states that there is a unique great circle through a given pair of nonantipodal points, and the second states that two distinct great circles intersect in a pair of antipodal points.

To see that this is a model for the projective plane, consider the correspondence between the "points" of the second and third model obtained by taking S as the center of the unit sphere and associating each ray through S with the "point" (that is, the pair of antipodal points) in which it meets the surface of the sphere. This correspondence is evidently one-one, is defined on every ray of the second model, and has every "point" of the third model in its range. Moreover, the image of a "line" under this mapping is a "line"; for a line in the second model is a plane through S, and its image under the map is the great circle in which this plane cuts the sphere. We can also put coordinates on the third model immediately: we associate with each "point" its ordinary coordinates in a coordinate system for space that has the center of the sphere as origin. Every "point," then, has two triples for coordinates —a triple (x_1, x_2, x_3) such that $x_1^2 + x_2^2 + x_3^2 = 1$, and the triple $(-x_1, -x_2, -x_3)$. The correspondence we have set up between the second and third models preserves coordinates, in the sense that homogeneous coordinates in the third system for a given point are also homogeneous coordinates for the corresponding point in the second model. The converse

is not quite true, but if (x_1, x_2, x_3) are homogeneous coordinates of a point M in the second model, then suitable multiples of them $(\lambda x_1, \lambda x_2, \lambda x_3)$ are homogeneous coordinates for the corresponding "point" in the third model. We need merely take λ so that $\lambda^2(x_1{}^2 + x_2{}^2 + x_3{}^2) = 1$.

As in the second model, a line has a homogeneous linear equation; conversely, the "points" whose coordinates satisfy a homogeneous linear equation lie on a "line." The linear equation, as in the second model, is the linear equation of the plane through S in which the given "line" (great circle) lies.

Fourth Model. The "points" in this model will be all the classes of proportional triples of real numbers, excluding the triple (0, 0, 0). This model is more abstract in one sense than the others. Instead of constructing more or less comprehensible geometric objects for our "points" and then showing how we can introduce coordinates in such a way that each point is assigned a complete class of proportional triples for its coordinates, we here simply state that the point is to *be* the class of proportional triples—we identify the point with its coordinates.

We define a line to be the set of all "points" for which any triple of coordinates satisfies some given linear homogeneous equation. It is then a purely algebraic fact that two distinct lines have a unique point in common and that there is a unique line through two given points. For example, the first of these states that there is a unique triple of real numbers (x_1, x_2, x_3), up to a factor of proportionality, that satisfies both of the equations:

$$Ax_1 + Bx_2 + Cx_3 = 0,$$

$$A'x_1 + B'x_2 + C'x_3 = 0,$$

where

$$A:B:C \neq A':B':C'.$$

This condition is what we need to ensure that the two lines are distinct.

The fourth model is perhaps the best, in that it keeps all the essential features, without any extraneous ones. If anything, all the other models, including the first, could best be called models of this one, in that they give concrete geometric realizations of it. When we define the real projective plane in this way, we can see immediately how to generalize it fruitfully—in one direction, by defining projective spaces of higher dimension, merely by considering classes of proportional n-tuples of real numbers (not all zero) as our points when we are defining the real projective $(n - 1)$-space, and in another direction, by allowing other fields than the field of real numbers. For example, we may define the *complex projective plane* to be the collection of all classes of proportional triples of *complex* numbers, not all zero. The definition of a line remains unchanged, except that we allow complex coefficients.

Instead of talking of models, we may use the important concept of isomorphism. We start by defining an *abstract projective plane* as follows:

A set P of objects called *points* is said to be a *projective plane* if there are defined in it certain subsets called *lines* such that:

(1) there is a unique line containing both of two distinct given points of P;

(2) there is a point common to any two lines;

(3) there exist at least four points, no three of which are collinear (that is, belong to a single line).

All our models of the real projective plane satisfy these axioms, but there are other objects which do so too (for example, the complex projective plane, whose definition we sketched earlier).

In definition (2), we do not need to specify a unique common point on two distinct lines, since this follows from (1).

Axiom (3) is included to rule out some trivial and uninteresting objects that would otherwise count as projective planes but have exceptional properties that make them a nuisance.

We define an *isomorphism* between two projective planes P, Q to be a one-one mapping of the projective plane P *onto* the projective plane Q such that the image of a line is a line.

Then we see that the four models of the real projective plane we have given are all isomorphic. We started by proving the first and fourth isomorphic, for this is really what we were doing when we introduced coordinates into the projective plane. We were indicating how to associate coordinates (that is, a point in the fourth model) with each point of the plane, and we proved that the image of a line was a line by showing that the equation of a line in the coordinate system we introduced was just what it should be for its image to be a line, according to the definition in the fourth model.

We next proved the second and first isomorphic, under the mapping that takes each point into the ray through S and that point, and, in the next section, showed independently that the second and fourth models were isomorphic. We then showed how the third model was isomorphic to both the second and the fourth. In terms of projective geometry, these four models are all the same thing, and we can use them as is most convenient to us. The fourth model is especially useful, because it allows us to use algebraic methods.

Isomorphism is a concept that occurs everywhere in mathematics, not only in projective geometry. To say that two objects in some branch of mathematics are isomorphic is to say that, as far as that branch of mathematics is concerned, they are the same. Thus, to say that our models of the real projective plane are isomorphic (with respect to the concepts of projective geometry, such as point, line, incident) is to say that in projective geometry they are all the same. In exactly the same way, Cartesian geometry may be thought of as Euclidean geometry carried out in an isomorphic copy of the Euclidean plane—the Cartesian plane. The isomorphism between them associates points with ordered pairs of real numbers, lines with the pairs that satisfy linear equations, and so on. All the concepts (such as parallelism, angle, area) of Euclidean geometry can be translated into the coordinate model, and, as far as Euclidean

geometry is concerned, the "pure" (synthetic) and coordinate versions are the same thing.

Two final points remain. We gave above a definition of an abstract projective plane and indicated that there are nonisomorphic objects, all of which are projective planes. We have given only one example of a projective plane (up to isomorphism) and will only be dealing with this one plane in this book, whatever model we may use for it. It is called the *real projective plane*, the meaning of *real* being that, in the fourth model, we take triples of *real* numbers.

The second point concerns the first model. We have talked of ideal points and the ideal line of a projective plane, but this term only arose in the first model. In the other models there are no points or lines that are distinguished in some intrinsic way from all the others. True, when we establish an isomorphism between the first model and another one, we can see what line and what points in the other model are made to correspond with the ideal ones. Thus, in the second model, the plane through S parallel to π is the ideal plane ("line") and the rays in it the ideal "points." But we could have established our isomorphism differently, by choosing a different plane π' out of which to construct Π. In this case the ideal points of the second model would be the rays parallel to π'. In the fourth model we take $x_3 = 0$ to be the ideal line, but we could equally well have rearranged our coordinates and found $x_2 = 0$ to be the special line. There is nothing special about the line $x_3 = 0$ *within* the fourth model. The definition of an abstract projective plane we gave says nothing about ideal points or lines.

The reason for talking of ideal points and lines when we are considering a projective plane is to indicate that it is to be thought of as constructed out of a certain given ordinary plane. Given a projective plane, we can always do this by choosing a line at random to be the ideal line, assigning coordinates so that this line has equation $x_3 = 0$, and then mapping the other points onto a Cartesian plane by sending the point (x_1, x_2, x_3) into the point (x, y), where $x = x_1/x_3$, $y = x_2/x_3$.

Thus, ideal points arise when we think of a projective plane in terms of affine space (for example, when we defined a perspectivity for a projective plane); when we study projective geometry as a "pure" subject, there are no ideal points.

2. Definition of a Projective Mapping

Definition. A *projective mapping* of a projective plane Π onto a projective plane Π' is a one-one mapping of Π onto Π' such that the images of three collinear points are collinear.

A projective mapping of Π onto itself is called a *projective transformation* of Π.

Just as for affine mappings, we can show that the inverse of a projective mapping is another and then that the collection of all projective transformations of Π forms a group.

The reader will notice that the definition we have given for a projective mapping is word for word the same as that which we gave for an affine mapping. However, these definitions refer to entirely different objects; an affine mapping is defined on an ordinary (affine) plane, while a projective mapping is defined on the projective plane. This fact makes these definitions the same only formally; actually, both the concept and the properties of projective mappings are distinct from those of affine mappings. There is a partial connection between them, which we can bring out as follows. Suppose that π and π' are two ordinary planes and Π and Π' the projective planes constructed from them in the usual way (first model). Let A be a projective mapping of Π onto Π'. Then we have

Theorem 1. *If A maps the ideal line of Π onto the ideal line of Π', then the restriction α of A to π is an affine mapping of π onto π'.*

Proof. Since A is one-one, so is α. For each point P' of π' there is a point P of Π such that $A(P) = P'$. Since A maps every ideal point onto an ideal point and P' is ordinary, P

must be ordinary, so that $\alpha(P) = P'$. Thus α is onto π'. It is defined on all π since A is. It remains to prove the line-preserving property. Let B, C, D be collinear points of π. Then they are also collinear points of Π. (That is, they lie on a projective line of Π; in fact, they lie on the projective line obtained by adding an ideal point to the ordinary line BCD of π.) So their images B', C', D' lie on a line L of Π'. But B', C', D' are ordinary points of Π', so they lie on the ordinary part l of L. But l is an ordinary line of π', so that B', C', D' are collinear points of π'. ▼

As an example illustrating this theorem, consider the perspectivity between two parallel planes π and π'. We showed how a perspectivity may be regarded as defined on the projective planes, and in this case the ideal line of Π is mapped onto the ideal line of Π'. The reader can check that the perspectivity does, in fact, define an affine mapping of π onto π'.

Theorem 2. *Let π and π' be affine planes, and let α be an affine mapping of π onto π'. Let Π and Π' be the projective planes constructed from π and π' in the usual way. Then there exists a unique projective mapping A of Π onto Π' whose restriction to π is α.*

A is called an *extension* of α, and the theorem states that an affine mapping of π has a unique projective extension to Π.

Proof. We know what effect A must have on the ordinary points of Π: just the effect of α. It remains to specify what A does to the ideal points of Π. Let M be a given ideal point of Π, and let L be any line through it. Let l' be the image under α of the ordinary part l of L, and let L' be the unique line of Π', the ordinary part of which is l'. Then we make A take M into the ideal point of L'.

Clearly, if A is to preserve collinearity on Π, it *must* be defined in this way. So we have already proved uniqueness

(if A does, in fact, do the trick); it remains to prove that A is a projective mapping.

We must check first that this is a good definition; that is, that we arrive at the same point M' when we make a different choice of L. Suppose that we had chosen L_0 to go through M instead. Then the ordinary part l_0 of L_0 is parallel to l. So its image l_0' under α is parallel to l' (since affine mappings preserve parallelism) and the ideal point of L_0' is the same as the ideal point M' of L'.

We need now to check that A is one-one. If M and N are distinct ordinary points, then $A(M) \neq A(N)$, since $A(M) = \alpha(M)$ and $A(N) = \alpha(N)$, and α is one-one. If one of M and N is ideal, and the other not, then one of the images is ideal and the other not, since A takes ideal points into ideal points and ordinary points into ordinary points. Suppose M and N are both ideal and that M_0 and N_0 are lines through them. Then M_0 and N_0 intersect in an ordinary point (since they only have one ideal point each, and these points are not the same, whereas any two projective lines intersect). The images under α of the ordinary parts m_0 and n_0 of M_0 and N_0 are intersecting ordinary lines m_0' and n_0'. So the projective lines of Π' whose ordinary parts are m_0' and n_0' have distinct ideal points, these points being the images of M and N under A.

We still need to show that A is onto and that it preserves collinearity. Like the arguments we have given so far, these are proved by routine arguing "from the definition," and we leave them to the reader. ▼

In view of these theorems, we will be able to use the language of affine transformations when we talk of projective transformations (see, for example, the second fundamental theorem in Section 3 below). For example, we may say that α *is a translation* of Π onto Π'. By this we mean that we are to think of Π and Π' as constructed from distinct planes π and π' of space and that α *induces a translation* of π onto π'. Similarly, we will find it convenient to be careless about the word

"perspectivity," meaning either the constructive perspectivity of two planes π and π' in space or the projective transformation determined by it on the projective spaces Π and Π' obtained from π and π' in the usual way. We first give a better definition of a perspectivity than the clumsy one we gave before (Section 1 of this volume).

Let π and π' be distinct affine planes in space, and let Π and Π' be their projective "completions." Let S be any point not on π or π'. Then we know how to map Π onto the projective space of rays through S. Let us call this mapping α. It is a projective mapping, as we may check from the definition of the mapping (Section 1.3 above) and of a projective map (this section). Let β similarly be the mapping of Π' onto the projective space of rays through S. Then $\beta^{-1}\alpha$ is a projective mapping of Π onto Π'. This mapping we call the *perspectivity* of Π onto Π' *with center S*.

Referring back to Fig. 2, we see clearly how the ideal points K' of Π' correspond to the points K of k in Π, via the ray through S and K, and how the points L' of l' in Π' have as inverse images the ideal points of Π, via the ray through S and L'.

Let us note that the line n of intersection of π and π' is pointwise invariant under the perspectivity and that the ideal point of this line (whether n is regarded as part of a projective line in Π or Π') is the same and is also invariant under the perspectivity. Note also that k and l' are both parallel to n, and so have the same ideal point. (We are using language loosely here; k and l' do not have ideal points since they are simply ordinary lines of π and π', but it is clear what we mean, and we shall not avoid such language in the future.) Unless π and π' are parallel, this is the only pair of corresponding points which are both special.

We leave it to the reader to show that the inverse mapping of a perspectivity (from Π to Π', say) is a perspectivity (from Π' to Π). We may thus talk of a perspectivity *between* two projective planes.

3. Two Fundamental Theorems on Projective Transformations

3.1. First Fundamental Theorem

We prove first a theorem that is important not only for projective geometry but also in practice; for example, in aerial photography (see 3.3 below). We emphasize that we are talking about the *real* projective plane: there exist projective planes for which the theorem is false.

Theorem 1. *Given any four points A, B, C, D of a projective plane Π, no three of which are collinear, and four points A', B', C', D' of a projective plane Π', no three of which are collinear, there exists one and only one projective mapping α of Π onto Π' that takes A, B, C, D into A', B', C', D', respectively.*

For our proof, we shall use the second model. The projective plane consisting of the rays through S (for the points) and the planes through S (for the lines) will be referred to as the projective "space" S. Moreover, the line SA will be referred to as the "point" A, and the plane through the "points" A and B (that is, the plane through S, A, B) will be referred to as the "line" AB.

Proof of Existence. Let S and S' be two "spaces," A, B, C, D four "points" of S, no three of which are "collinear," and A', B', C', D' four "points" of S', no three of which are "collinear."

Let E be any point on the "line" D (other than S), and let the planes through E parallel to the "lines" BC, CA, AB meet the "points" A, B, C, respectively, in the points E_1, E_2, E_3 (Fig 5).

Define E', and E_1', E_2', E_3' in precisely the corresponding manner for S'.

Consider the affine transformation α^* of space taking S, E_1, E_2, E_3 into S', E_1', E_2', E_3', respectively (see Volume 1,

Section 32). Then an easy argument shows that this transformation takes E into E' and takes the "points" A, B, C, D into the "points" A', B', C', D'.

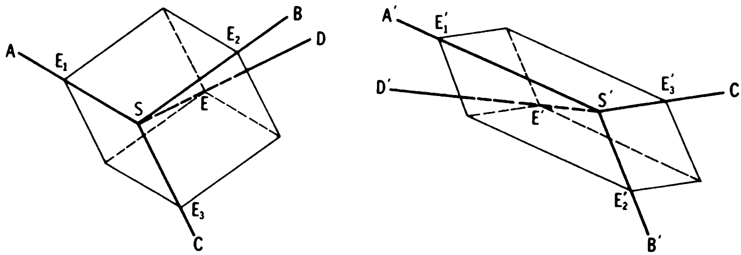

Fig. 5

Since α^* maps lines through S onto lines through S' and maps planes through S onto planes through S', it induces a mapping α of the "space" S onto the "space" S' in the obvious way. This mapping α is "line" preserving, is one-one (a fortiori, since α^* is) and is onto (again because α^* is). Thus it is a projective mapping of one "space" onto the other, and we have already seen that it takes the "points" A, B, C, D into the "points" A', B', C', D', respectively. ▼

Proof of Uniqueness. Let α be any projective mapping of the "space" S onto the "space" S', under which the "points" A, B, C, D go into A', B', C', D', respectively.

Let E be the "point" in which the "lines" AD and BC meet and E' be the "point" in which the "lines" $A'D'$ and $B'C'$ meet (Fig. 6).

Since the image under α of a "line" is a "line" and the properties of belonging (of "point" to "line") and of intersection (of "lines") are preserved under projective mappings, the "point" E' must correspond under α to E. The "points" C, D, E are not "collinear," and none of them lies in the "line" AB; so the corresponding assertions are true of the "points" C', D', E'. Let π and π' be planes not passing through S or S' and parallel to the "lines" AB and $A'B'$, respectively.

Let c, d, e be the points of intersection of π with the "points" C, D, E, respectively, and c', d', e' the points of intersection of π' with C', D', E', respectively.

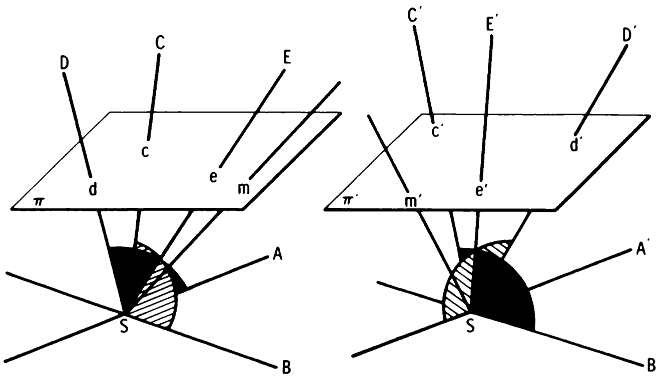

Fig. 6

Let m be any point of π; then the "point" Sm does not lie on the "line" AB, and therefore its image $S'm'$ does not lie on $A'B'$, and it meets π' in some point m'. Let β be the correspondence between π and π' in which each point m goes into m'. Then we assert that β is an affine mapping. We need to check the following:

(1) β is well defined on every point of the (ordinary) plane π;

(2) the points m_1', m_2' corresponding to distinct points m_1, m_2 of π are distinct;

(3) any point m' of π' has an inverse image m in π;

(4) the images m_1', m_2', m_3' of three collinear points m_1, m_2, m_3 are collinear.

We leave the straightforward verification to the reader.

Suppose now that there is another projective mapping α' of the "space" S onto the "space" S', which also sends A, B, C, D into A', B', C', D', respectively. Then the affine mapping β' induced by it on π, in the same way as β is induced by α,

has the same effect as β on the noncollinear points c, d, e. So (Volume 1, Section 25). $\beta' = \beta$.

It follows that any "point" of S not lying on AB has the same image under α and α'. It remains to show that they coincide also on AB

Now the "space" S excluding the "points" of AB is precisely an ordinary affine space. For we obtain affine space by deleting all the points of the special line from projective space and defining what is left of the projective lines to be our lines, the remaining points to be our points, and lines to be parallel if their projective analogs intersected in a special point (so that now they do not intersect). By taking AB to be the ideal "line" of the "space" S and $A'B'$ the ideal "line" of space S', we are left, after removal of these "lines", with affine spaces on which α and α' are affine mappings that agree on every point, that is, are identical. But we proved in the last section that an affine mapping has a unique extension to a projective mapping (on the projective space constructed out of the affine one). So the extensions of α and α' to all of the space S are the same, that is, $\alpha = \alpha'$. ▼

If the reader thinks that this argument is a little abstract, he can supply a direct one merely by imitating the argument of the proof of Theorem 2, Section 2 above.

Corollary. *A projective transformation of the plane that fixes four points, no three of which are collinear, is the identity.*

For, if the points are A, B, C, D, then by the theorem there exists a unique projective mapping that takes them into A, B, C, D, respectively; but the identity is such a mapping.

3.2. Second Fundamental Theorem

We now show that any projective mapping reduces either to an affine mapping or to the successive operations of moving an affine plane and taking a perspectivity.

Theorem 2. *Let Π and Π' be projective planes obtained in the usual way from the planes π and π' of space, and let α be a projective mapping of Π onto Π'. Then either α is induced by an affine mapping of π onto π' or $\alpha = \beta\gamma$, where γ is an orthogonal mapping of Π to some projective plane Π^*, and β is the perspectivity from a suitable center S of Π^* onto Π'.*

See the discussion in Section 2 above.

Proof. Case 1. The ideal line of Π is the image under α of the ideal line of Π'. We have already dealt with this case (Theorems 1 and 2, Section 2 above), and know that the restriction β of α onto π is an affine mapping of π onto π' and that α is the unique extension of β to a projective mapping of Π onto Π'.

Case 2. The ideal line of Π' is the image under α of the line d of π (together with the ideal point on d).

Consider in π any square $ABCD$ whose vertices C and D lie on d. Since A and B do not lie on d, their images A' and B' under α will be ordinary points of π' (Fig. 7). Suppose that

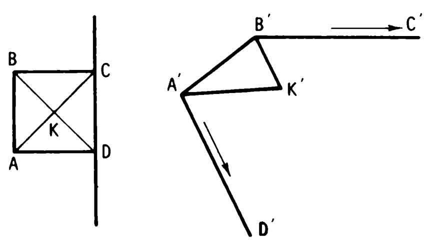

Fig. 7

the diagonals AC and BD meet in K. Then K does not lie on d, so that its image K' under α is an ordinary point of π'.

Since C and D lie on d, their images C' and D' under α will be special points of Π', and, in π', $A'K'$ is parallel to $B'C'$, and $B'K'$ to $A'D'$.

Let π_0 be any plane through $A'B'$ (Fig. 8). Choose points C_0 and D_0 in π_0 so that $A'B'C_0D_0$ is a square. Draw the parallel through C_0 to $B'K'$, and the parallel through D_0 to $A'K'$. These lines are parallel to the plane π' and are not parallel to each other (since $B'K'$ and $A'K'$ are not parallel), so that they meet in some point S. We take S to be the center of our perspectivity β. It is clear that β is a perspectivity of Π_0 onto Π' under which A', B', C_0, D_0 go into A', B', C', D', respectively.

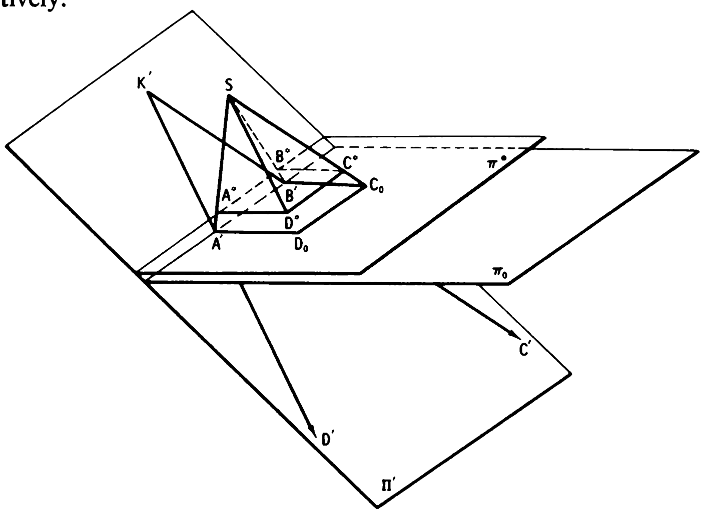

Fig. 8

It is clear that every plane parallel to π_0 cuts the pyramid $SA'B'C_0D_0$ in a square. Choose π^* parallel to π_0 so that the square $A^*B^*C^*D^*$ of intersection is congruent to $ABCD$. Let γ be an orthogonal transformation of space taking A, B, C to A^*, B^*, C^* respectively. We may regard γ as an orthogonal transformation of Π onto Π^*.

It is clear that the perspectivity of Π^* onto Π' with center S takes A^*, B^*, C^*, D^* into A', B', C', D', respectively, so that $\beta\gamma$, like α, takes A, B, C, D into A', B', C', D', respectively.

Since no three of the points A, B, C, D are collinear, and no

three of A', B', C', D' are collinear, there is a unique projective transformation taking the one quadruple of points into the other. So we must have $\alpha = \beta\gamma$. ▼

We have thus completed the proof of our theorem.

3.3. Applications to Aerial Photography

We now consider an application of the two fundamental theorems to the "righting" of skew photographs. Suppose we are carrying out a photographic survey of a stretch of land from an airplane. If, at the moment a shot is taken, the axis of the camera is vertical, the representation of a stretch of flat country on the film is similar (in the strict mathematical sense) to the country itself. We leave the simple proof of this to the reader. In actual fact, because of the inevitable slight rolling and pitching of the aircraft, the axis of the camera changes direction from moment to moment, and the film will have various perspective mappings of the ground (the center of perspectivity is the lens, and the planes are the plane of the ground and the plane of the film).

However, successive photographs are taken at such short intervals that each successive shot shows part of the ground covered by the preceding shot. This fact gives us a method of correcting for the distortions in the photographs introduced by the tilt of the airplane: the corrections are made by means of a simple instrument called a photographic transformer. We choose four recognizable points on the negative, no three of which are collinear, and find their positions on a map. Above the negative there is a point source of light, and the negative is moved about until the rays through the four points exactly hit the corresponding points on the map (or on a special piece of paper on which the relative positions of the four objects are marked). The negative is fixed in this position, and the map or paper is replaced by photographic paper. The light source then produces a positive image on the photographic paper, which exactly represents the ground. Having done this to every photograph, we can stick the photographs together along

corresponding boundaries to get an extended and faithful picture of the land. Note that if the work is done accurately, we need only to know four points on the map identifiable on one of the pictures (say the first), because the corrected positive of the first picture can serve as the "map" to align the second, the corrected second photograph to align the third, and so on. We can do this because, as we said above, part of the ground covered by each photograph after the first has already been covered by its predecessor.

4. Cross Ratio

4.1. The Cross Ratio of Four Points on a Line

One of the more important concepts of real projective geometry is the *cross ratio* of four collinear points.

We shall introduce this concept on the first model of the projective plane and then extend it to the second also.

Let Π be the projective plane obtained in the usual way from the Euclidean plane π. Let A, B, C, D be four collinear points of Π, and suppose first that these points are distinct and not special. In this case we define the cross ratio $(ABCD)$ of the four points (taken in that order) to be the real number

$$\frac{\overrightarrow{AC}}{\overrightarrow{CB}} : \frac{\overrightarrow{AD}}{\overrightarrow{DB}}.$$

Let us consider what happens to this cross ratio when we keep three of the points, A, B, C, say, fixed and let the fourth point D tend to infinity along the line, that is to say, let the length AD tend to infinity. We find that the value of the cross ratio $(ABCD)$ tends to a limit:

$$\lim_{D\to\infty}(ABCD) = \lim_{AD\to\infty} \frac{\overrightarrow{AC}}{\overrightarrow{CB}} : \frac{\overrightarrow{AD}}{\overrightarrow{DB}}$$

$$= \frac{\overrightarrow{AC}}{\overrightarrow{CB}} : \lim_{AD\to\infty} \frac{\overrightarrow{AD}}{\overrightarrow{DB}} = -\frac{\overrightarrow{AC}}{\overrightarrow{CB}}.$$

In exactly the same way, we find that

$$\lim_{C\to\infty}(ABCD) = -\frac{\overrightarrow{BD}}{\overrightarrow{DA}},$$

$$\lim_{B\to\infty}(ABCD) = -\frac{\overrightarrow{CA}}{\overrightarrow{AD}},$$

$$\lim_{A\to\infty}(ABCD) = -\frac{\overrightarrow{DB}}{\overrightarrow{BC}}.$$

In the case where one of the four points A, B, C, D is the ideal point on the line containing the other three, we take the cross ratio of the four to be the limit of the values of the cross ratios of four points, three of which are the given ordinary points, as the fourth tends to infinity. Thus, if the ideal point of the four is D, then, by definition,

$$(ABCD) = -\frac{\overrightarrow{AC}}{\overrightarrow{CB}};$$

if C, then

$$(ABCD) = -\frac{\overrightarrow{BD}}{\overrightarrow{DA}};$$

if B, then

$$(ABCD) = -\frac{\overrightarrow{CA}}{\overrightarrow{AD}};$$

and, finally, if the ideal point is A, we take

$$(ABCD) = -\frac{\overrightarrow{DB}}{\overrightarrow{BC}}.$$

We see that in each of the four cases, the cross ratio reduces to the ordinary ratio of two of the segments determined by the remaining three points. We have not yet defined the cross ratio when all four points lie on the ideal line. We will come back to this later.

4.2. The Cross Ratio of Four Lines of a Sheaf

Let S be an ordinary point of Π, and let a, b, c, d be four distinct lines through it. Let A, B, C, D be points on the respective lines, distinct from S (Fig. 9). We define the cross ratio of the four lines a, b, c, d to be the real number

$$(abcd) = \frac{[SAC]}{[SCB]} : \frac{[SAD]}{[SDB]},$$

where $[SAC]$, for example, denotes the oriented area of the oriented triangle SAC.

This definition is independent of the choice of A, B, C, D. We prove, for example, that it is independent of D.

Suppose that D' is a point of d distinct from D or S. We distinguish two cases:

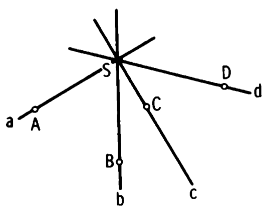

Fig. 9

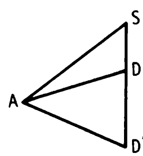

Fig. 10

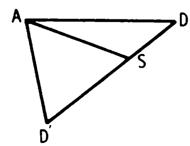

Fig. 11

Case 1. D and D' lie on the same side of S (Fig. 10). In this case the triangles SAD and SAD' have the same orientation, and so do the triangles SDB and $SD'B$. So the ratios

$$\frac{[SAD]}{[SDB]} \quad \text{and} \quad \frac{[SAD']}{[SD'B]}$$

have the same sign. We show that their absolute values are also the same:

$$\left|\frac{[SAD]}{[SDB]}\right| = \frac{\frac{1}{2}SA \cdot SD \cdot \sin \angle ASD}{\frac{1}{2}SB \cdot SD \cdot \sin \angle BSD} = \frac{SA \cdot \sin \angle ASD}{SB \cdot \sin \angle BSD};$$

$$\left|\frac{[SAD']}{[SD'B]}\right| = \frac{\frac{1}{2}SA \cdot SD' \sin \angle ASD'}{\frac{1}{2}SB \cdot SD' \sin \angle BSD'} = \frac{SA \sin \angle ASD}{SB \sin \angle BSD}.$$

Case 2. D and D' lie on opposite sides of S (Fig. 11). In this case, the triangles SAD and SAD' have opposite orientations and so do the triangles SDB and $SD'B$. So the ratios

$$\frac{[SAD]}{[SDB]} \quad \text{and} \quad \frac{[SAD']}{[SD'B]},$$

once again, have the same sign. The absolute values of these ratios are also equal, the proof being word for word as before.

We postpone the definition of the cross ratio $(abcd)$ in the case where S is an ideal point.

We now show how the concept of cross ratio for points on a line and cross ratio for lines through a point are intertwined. Let us say that four points A, B, C, D of a line are *in perspective* with four lines a, b, c, d of a sheaf if each line passes through the corresponding point. We shall further say that the four collinear points A, B, C, D are in perspective with the four collinear points A', B', C', D' if the quadruples lie on different lines and the lines AA', BB', CC', DD' intersect in some point S. Note that we are saying that two quadruples of collinear points are in perspective if they are each in perspective with the same quadruple of lines of a sheaf, in this case the quadruple $(abcd)$ through S, where $a = AA'$, and so forth. Finally, we shall say that two quadruples $(abcd)$ and $(a'b'c'd')$ of sheaves with distinct centers S, S' are in perspective if they are each in perspective with the same quadruple $(ABCD)$ of collinear points, that is, if the points of intersection of corresponding lines are collinear. We do not exclude the case where one pair of lines, a and a', say, coincide. Similarly, we allow two quadruples of points to be in perspective even if $A = A'$ in the notation above. In the one case, we regard the point of intersection of a and a' as the point of intersection of a with the line BCD, and, in the other, we take the line a joining A and A' to be the line AS, where S is the point of intersection of $b = BB'$, c and d. We advise the reader to sketch a figure.

Theorem 1 (provisional). *The cross ratio of four points of an ordinary line is equal to the cross ratio of any quadruple of lines of a sheaf that is in perspective with them.*

Proof. Case 1. All four points A, B, C, D are ordinary.

$$(abcd) = \frac{[SAC]}{[SCB]} : \frac{[SAD]}{[SDB]}.$$

The triangles SAC and SCB have the same height. If they have the same orientation, then the ratio $[SAC]$: $[SCB]$ is positive and equal to the ratio of the bases of the triangles:

$$\frac{[SAC]}{[SCB]} = \frac{AC}{CB} = \frac{\overrightarrow{AC}}{\overrightarrow{CB}},$$

since in this case the vectors $\overrightarrow{AC}$ and $\overrightarrow{CB}$ have the same direction.

If the triangles have opposite orientation, the ratio of their oriented areas is negative, and its absolute value is the ratio of their bases:

$$\frac{[SAC]}{[SCB]} = -\frac{AC}{CB} = \frac{\overrightarrow{AC}}{\overrightarrow{CB}},$$

since the vetcors $\overrightarrow{AC}$ and $\overrightarrow{CB}$ have opposite directions.

Thus,

$$(abcd) = \frac{[SAC]}{[SCB]} : \frac{[SAD]}{[SDB]} = \frac{\overrightarrow{AC}}{\overrightarrow{CB}} : \frac{\overrightarrow{AD}}{\overrightarrow{DB}} = (ABCD).$$

Case 2. One of the points A, B, C, D is ideal. Suppose, for example, that D is ideal. Then d is parallel to ABC. Let D^* be any ordinary point on d; by definition, we have

$$(abcd) = \frac{[SAC]}{[SCB]} : \frac{[SAD^*]}{[SD^*B]}$$

$$= \frac{\overrightarrow{AC}}{\overrightarrow{CB}} : -\frac{[ASD^*]}{[BSD^*]} = -\frac{\overrightarrow{AC}}{\overrightarrow{CB}} = (ABCD).$$

We may show by similar arguments that

$$(abcd) = (ABCD),$$

whichever of the points A, B, C, D is ideal. ▼

Corollary 1. *If four lines a, b, c, d of an ordinary sheaf are cut by ordinary lines l and l' in A, B, C, D and A', B', C', D', respectively, then*

$$(ABCD) = (A'B'C'D').$$

Proof. An ordinary sheaf means, of course, a sheaf whose vertex S is an ordinary point. By the theorem we have just proved,

$$(ABCD) = (abcd),$$

and

$$(A'B'C'D') = (abcd).$$

Thus

$$(ABCD) = (A'B'C'D'). \quad \blacktriangledown$$

Corollary 2. *If the quadruples $(abcd)$ and $(a'b'c'd')$ of lines from distinct sheaves are in perspective, then*

$$(abcd) = (a'b'c'd').$$

Proof. Each of the cross ratios is equal to $(ABCD)$, where A is the point of intersection of a and a', and so on. $\blacktriangledown$

We are now in a position to complete the definition of cross ratio. We have not yet defined the cross ratio of four points all of which are ideal, or of four lines of a sheaf through an ideal point.

We say that the cross ratio $(ABCD)$ of four ideal points is the cross ratio $(abcd)$, where $a = SA$, and so on, and S is some fixed ordinary point. This definition is independent of our choice of S, for if $a' = S'A$, and so on, then a' is parallel to a, and so on, and $(a'b'c'd') = (abcd)$.

We say that the cross ratio $(abcd)$ of four lines through an ideal point is the cross ratio $(ABCD)$, where A is the point of intersection of l and a, and so on, and l is some fixed ordinary line. This definition is independent of our choice of l, for if A' is the point of intersection of l' and a, and so on, then even the ordinary ratios

$$\frac{\overrightarrow{AC}}{\overrightarrow{CB}} \quad \text{and} \quad \frac{\overrightarrow{AD}}{\overrightarrow{DB}}$$

will not change;

$$\frac{\overrightarrow{AC}}{\overrightarrow{CB}} = \frac{\overrightarrow{A'C'}}{\overrightarrow{C'B'}}, \qquad \frac{\overrightarrow{AD}}{\overrightarrow{DB}} = \frac{\overrightarrow{A'D'}}{\overrightarrow{D'B'}},$$

and, therefore,

$$(ABCD) = (A'B'C'D').$$

We are now in a position to restate the theorem of this section and its two corollaries without any qualification.

Theorem I (final). *The cross ratio of four collinear points is equal to the cross ratio of any four lines of a sheaf in perspective with them.*

If four points of one line are in perspective with four points of another, then the corresponding cross ratios are equal.

If four lines of one sheaf are in perspective with four lines of another, the corresponding cross ratios are equal.

4.3. Cross Ratio in the Second Model

We now define the cross ratio of four "points" of a "line" and four "lines" of a "sheaf" in the second model. For the sake of symmetry, instead of talking of "lines" of a "sheaf," we shall talk of "lines" through a "point."

We recollect that the "points" are the lines of a sheaf with vertex S in space, and the "lines" are the planes through S. The condition that four "points" should lie on a "line" is that the four corresponding rays through S should be coplanar. It is therefore natural, in view of the last section, to define the cross ratio $(ABCD)$ of four "points" as follows: on the "points" A, B, C, D choose points K, L, M, N, respectively, and set

$$(ABCD) = \frac{[SKM]}{[SML]} : \frac{[SKN]}{[SNL]}$$

with the notation as before. The value obtained for $(ABCD)$ is independent of our choice of K, L, M, N, as was shown in Section 4.2 above.

We define the cross ratio of four "lines" through a "point"

to be the cross ratio of the four "points" of intersection of these "lines" and some other "line" not passing through the same "point" (vertex) as the given four. To show that this is independent of our choice of auxiliary "line", we proceed as follows. Suppose the given "lines" a, b, c, d are the planes SAA', and so on, shown in Fig. 12. The "point" of intersection of these "lines" is the vertical line in this figure. Let A, B, C, D be the "points" of intersection of a, b, c, d, with the "line" l (the horizontal plane $SKLMN$ in the figure). Let l' be another "line" meeting a, b, c, d in the "points" K', L', M', N', respectively (l' is the plane $SK'L'M'N'$). Let K, L, M, N be any four points on A, B, C, D, respectively, and let their parallel projections onto l' and parallel to the vertical axis (the "point" of intersection of a, b, c, d) be K', L', M', N', respectively. Since under parallel projection ratios of oriented areas are unaltered, we have

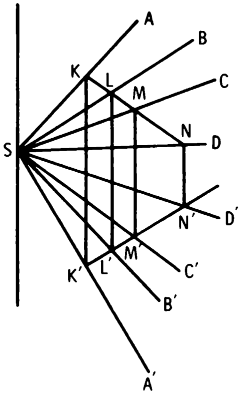

Fig. 12

$$\frac{[SKM]}{[SML]} : \frac{[SKN]}{[SNL]} = \frac{[SK'M']}{[SM'L']} : \frac{[SK'N']}{[SN'L']}.$$

But the left side is the value for the cross ratio $(abcd)$ that we obtain when we use l as the auxiliary "line," and the right side is the value for $(abcd)$ that we obtain when we use l' as our auxiliary "line." We have thus shown that our definition is independent of the choice of an auxiliary "line."

We now show that, under the correspondence between the first and second models of the projective plane (Section 1.3 above), cross ratios are preserved. That is to say, the cross ratio of four collinear points in the first model is equal to the cross ratio (as defined for the second model) of the four corresponding "points" of the second model. The latter will, of course, automatically be "collinear."

If the line on which the points A, B, C, D lie is an ordinary one, and the corresponding "points" in the second model are A^*, B^*, C^*, D^* (that is, A^* is the line SA, and so on,) then this follows from the definition of the cross ratio (A, B, C, D) and the results of Section 4.2. If A, B, C, D are ideal points, let a^*, b^*, c^*, d^* be four lines in π, one each from the four sheaves to which A, B, C, D, respectively, correspond. Then A^*, B^*, C^*, D^* will be lines through S parallel to a^*, b^*, c^*, d^*, respectively. So

$$(A^*B^*C^*D^*) = (a^*b^*c^*d^*).$$

But

$$(ABCD) = (a^*b^*c^*d^*)$$

(by the definition of the cross ratio of four ideal points), so that

$$(ABCD) = (A^*B^*C^*D^*).$$

The following theorem is also true:

Theorem 2. *The cross ratio of four lines through a point in the first model is equal to the cross ratio (as defined in the second model) of the four corresponding "lines" (all passing through the corresponding "point") in the second model.*

The proof is similar to the one above and is left to the reader.

4.4. Invariance of Cross Ratio Under Projective Mappings

Theorem 3. *Cross ratio is an invariant of projective mappings. This means that if A, B, C, D are four collinear points of a projective plane Π and α is a projective mapping of Π onto a projective plane Π' taking these points into A', B', C', D', respectively, then $(A, B, C, D) = (A', B', C', D')$.*

Proof. We use Theorem 2, Section 3 above.

Case 1. α is induced by an affine transformation of π onto π'. If A, B, C, D are ordinary, then the relation $(ABCD) = (A'B'C'D')$ follows from the fact that affine transformations preserve even the ratios of collinear segments, all the more so their cross ratios. If D, say, is ideal, then $(ABCD) = \lim (ABCD^*)$ as D^* recedes to infinity along the line ABC. But $(ABCD^*) = (A'B'C'D^{*\prime})$, and as D^* goes to infinity along ABC, $D^{*\prime}$ goes to infinity along $A'B'C'$. So

$$(ABCD) = \lim_{D^* \to \infty} (ABCD^*) = \lim_{D^{*\prime} \to \infty} (A'B'C'D^{*\prime}) = (A'B'C'D').$$

If A, B, C, D are all special, let us choose points A_0, B_0, C_0, D_0 of π that are in perspective with A, B, C, D. If A_0' is the image under α of A_0, and so on, then A', B', C', D' are in perspective with A_0', B_0', C_0', D_0', so that by what we have just proved and Theorem 1 of Section 4.2,

$$(ABCD) = (A_0B_0C_0D_0) = (A_0'B_0'C_0'D_0') = (A'B'C'D').$$

Case 2. $\alpha = \beta\gamma$ is the product of a rigid motion of π (or rather the mapping induced on Π by such a map of π) and a perspectivity. Under the rigid motion γ of π, cross ratios are preserved, inasmuch as both lengths and areas are preserved. That β preserves cross ratios follows from Theorem 1 in Section 4.2. ▼

We thus see a steady process of generalization. Under orthogonal mappings, lengths are preserved; under similarity mappings, lengths, in general, are not preserved, but ratios of lengths are; under affine mappings, ratios are not, in general, preserved, but ratios of segments on the same line are; under projective mappings, ratios are not, in general, preserved (when we think of the mappings confined to the ordinary part of the plane) even for segments on the same line, but ratios of ratios of adjacent segments on a line are. The fact that affine mappings preserve ratios of adjacent segments on a line may be regarded as a special case of this theorem. For let α be an affine mapping of π onto π', and let A be the induced mapping

of Π onto Π'. Let P, Q, R be three collinear points and P', Q', R' their images under α. Let X be the ideal point of the line PQR. Then A takes P, Q, R into P', Q', R', respectively, and X into the ideal point X' of $P'Q'R'$. So, by the theorem,

$$(PRQX) = (P'R'Q'X');$$

that is,

$$PQ : QR = P'Q' : Q'R'.$$

5. Harmonic Sets

5.1. Definitions and Examples

We say that four collinear points A, B, C, D of a projective plane form a *harmonic set* (or *are harmonic*) provided their cross ratio has the value -1:

$$(ABCD) = -1.$$

We say equivalently that D is the *fourth harmonic* to A, B, C. This expression is justified by the fact that given A, B, C, there is one and only one point D such that $(ABCD) = -1$. For a proof, use Example 7 below.

We say similarly that four lines a, b, c, d through a point form a *harmonic set*, or that d is *the fourth harmonic* to a, b, c, if we have

$$(abcd) = -1.$$

Let us consider some examples:

Example 1. If A and B are ordinary points, and C is the midpoint of the segment AB, then the fourth harmonic to A, B, C is the ideal point on AB. For, in this case (see Section 4.1 above),

$$(ABCD) = -\frac{\overrightarrow{AC}}{\overrightarrow{CB}} = -1.$$

Conversely, if $(ABCD) = -1$ and D is the ideal point on AB, then C is the midpoint of AB, since, from

$$(ABCD) = -\frac{AC}{CB} = -1,$$

it follows that $AC = CB$, which is precisely to say that C is the midpoint of AB.

Example 2. Let A, B, C, D be ordinary points that form a harmonic set. Let S be any ordinary point of the projective plane and l any ordinary line, not through S, and parallel to SD. Suppose the lines through S and A, B, C meet l in A', B', C', respectively. The C' is the midpoint of $A'B'$.

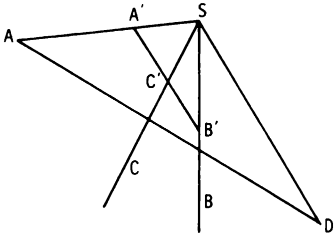

Fig. 13

For

$$-1 = (ABCD) = (A'B'C'D'),$$

where D' is the ideal point on l. It now follows, from Example 1, that C' is the midpoint of $A'B'$ (Fig. 13).

Example 3. Let a and b be nonparallel lines, and let c and d be their angle bisectors. Then the four lines are harmonic:

$$(abcd) = -1.$$

For suppose a parallel l to d meets a, b, c, d in A, B, C, D, respectively (in particular, D is the ideal point of l). Then $AC = CB$, so that $(ABCD) = -1$, by example 1, so that $(abcd) = -1$, by Theorem 1 of Section 4.2 (Fig. 14).

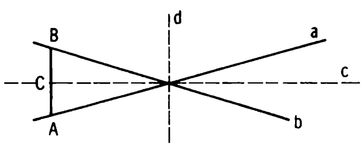

Fig. 14

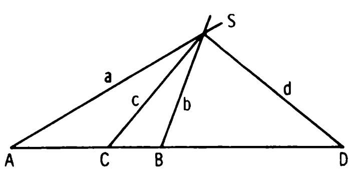

Fig. 15

We may establish this fact another way. Let l' be a line meeting a, b, c, d in the ordinary points A, B, C, D, respectively. Then, by a theorem on the internal and external angle bisectors of a triangle (Fig. 15),

$$\frac{\overrightarrow{AC}}{\overrightarrow{CB}} = \frac{AC}{CB} = \frac{AS}{SB},$$

$$-\frac{\overrightarrow{AD}}{\overrightarrow{DB}} = \frac{AD}{DB} = \frac{AS}{SB},$$

whence

$$(ABCD) = \frac{\overrightarrow{AC}}{\overrightarrow{CB}} : \frac{\overrightarrow{AD}}{\overrightarrow{DB}} = -1.$$

Example 4. If the four ordinary points A, B, C, D are harmonic, then the midpoint of CD lies outside the segment AB. If C divides AB in the ratio λ, then the midpoint H of CD divides AB in the ratio $-\lambda^2$.

For suppose that we have chosen an origin and a unit length on the line $ABCD$ and that the coordinates of A, B, C, D are x_1, x_2, x_3, x_4, respectively. Then, since

$$\frac{\overrightarrow{AC}}{\overrightarrow{CB}} = \lambda$$

and

$$\frac{\overrightarrow{AC}}{\overrightarrow{CB}} : \frac{\overrightarrow{AD}}{\overrightarrow{DB}} = -1,$$

we find

$$\frac{\overrightarrow{AD}}{\overrightarrow{DB}} = -\lambda,$$

so that D divides AB in the ratio $-\lambda$. We have shown, incidentally, that one of C and D lies within AB and the other outside. We can now calculate the coordinates of C and D in terms of λ:

$$x_3 = \frac{x_1 + \lambda x_2}{1 + \lambda},$$

$$x_4 = \frac{x_1 - \lambda x_2}{1 - \lambda},$$

so that the coordinate x_5 of the midpoint H of CD is

$$x_5 = \frac{x_3 + x_4}{2} = \frac{x_1 - \lambda^2 x_2}{1 - \lambda^2},$$

which is precisely to say that H divides AB in the ratio $-\lambda^2$.

Example 5. Let ASB be a triangle with a right angle at S (Fig. 16). Let SC be an altitude, and let SK and SL be the angle bisectors of the triangle at the vertex S. We assume that $SA \neq SB$,

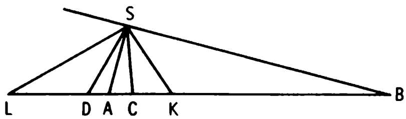

Fig. 16

so that L is an ordinary point and C and K do not coincide. Let SD be a median of the triangle LSK. Then D is the fourth harmonic to A, B, C.

For

$$\frac{\overrightarrow{AK}}{\overrightarrow{KB}} = \frac{AS}{SB} = \lambda,$$

$$\frac{\overrightarrow{AL}}{\overrightarrow{LB}} = -\frac{AS}{SB} = -\lambda,$$

$$\frac{\overrightarrow{AC}}{\overrightarrow{CB}} = \frac{AS^2}{BS^2} = \lambda^2$$

Since D is the midpoint of KL,

$$\frac{\overrightarrow{AD}}{\overrightarrow{DB}} = -\lambda^2,$$

so that

$$(ABCD) = -1.$$

Example 6. Let A, B, C, D be a harmonic set of ordinary points and let O be the midpoint of AB. Then

$$OC \cdot OD = OA^2.$$

For let us choose an origin at O and a unit of length. Then we may set

$$O(0), \quad A(a), \quad B(-a), \quad C(c), \quad D(d),$$

so that

$$(ABCD) = \frac{\overrightarrow{AC}}{\overrightarrow{CB}} : \frac{\overrightarrow{AD}}{\overrightarrow{DB}} = \frac{c-a}{-a-c} : \frac{d-a}{-a-d} = -1,$$

whence

$$a^2 = cd.$$

Example 7. If A, B, C, D are ordinary points forming a harmonic set,

$$\frac{1}{\rho} = \frac{1}{2}\left(\frac{1}{\rho_1} + \frac{1}{\rho_2}\right),$$

where $\rho = AB$, $\rho_1 = AC$, $\rho_2 = AD$ are the directed lengths (Fig. 17).

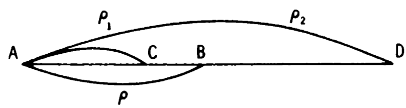

Fig. 17

For we have

$$(ABCD) = \frac{\overrightarrow{AC}}{\overrightarrow{CB}} : \frac{\overrightarrow{AD}}{\overrightarrow{DB}} = \frac{\rho_1}{\rho - \rho_1} : \frac{\rho_2}{\rho - \rho_2} = -1,$$

whence

$$\frac{\rho_1}{\rho - \rho_1} = \frac{\rho_2}{\rho - \rho_2}$$

or

$$\frac{1}{\rho} = \frac{1}{2}\left(\frac{1}{\rho_1} + \frac{1}{\rho_2}\right).$$

The reader may check that the result still holds when one of the points is ideal, provided that we assign it coordinate ∞ (with origin at A) and that we define $1/\infty = 0$.

Example 8. Let a and b be the diagonals of a parallelogram and c, d the parallels to its sides through its center. Then a, b, c, d form a harmonic set (Fig. 18).

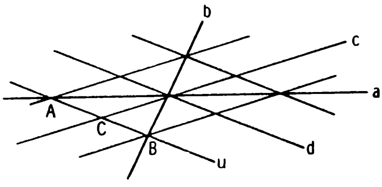

Fig. 18

For let u be a side of the parallelogram parallel to d, and let a, b, c, d meet u in A, B, C, D, respectively (D will be the special point of u). Then C is the midpoint of AB, and the result follows from Example 1.

Example 9. If D is the fourth harmonic of A, B, C, then C is the fourth harmonic of A, B, D. Thus, given A and B, C and D play symmetric roles, and we call each the *harmonic conjugate* of the other (with respect to A and B). Or we simply say C and D *are harmonic conjugates* with respect to A and B.

Suppose first that the four points are ordinary. Then

$$(ABDC) = \frac{\overrightarrow{AD}}{\overrightarrow{DB}} : \frac{\overrightarrow{AC}}{\overrightarrow{CB}} = \frac{1}{(ABCD)} = -1.$$

If one of the points is ideal, we leave the proof to the reader, and if all four are, we can use the argument from Case 1 of Theorem 3, Section 4.4 of this volume.

Example 10. If A, B, C, D are harmonic, then A and B are harmonic conjugates with respect to C and D. We leave the proof to the reader.

5.2. Methods of Constructing the Fourth Harmonic to Three Given Points

There are various methods of constructing the harmonic conjugate D of a point C with respect to two given points A and B. We give three.

1. Given the collinear points A, B, C, suppose first that one of the points is ideal. If C is ideal, D is the midpoint of AB (Example 1 above), and the construction is easy. If A or B is ideal, we mark off BD or AD, respectively equal to BC (or AC) and on the other side.

Suppose now that A, B, C are all ordinary. Draw any line through B and mark off equal segments BC' and $C'A'$ on it. Let S be the point (ordinary or ideal) of intersection of AA' and CC'. The parallel through S to BA' meets AB in the required fourth harmonic D. If S is ideal, we must interpret the line SD to be the line passing through the ideal points on both AA' (or CC') and $A'B$, so that it is the ideal line and meets AB in the ideal point D of AB. This happens (as the reader may verify) only when $AC = CB$ (Fig. 19).

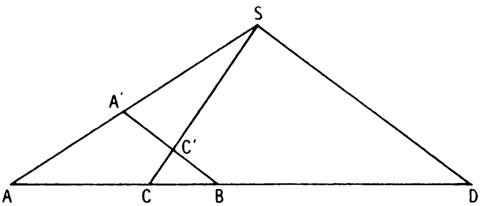

Fig. 19

2. Suppose that we are given the collinear ordinary points A, B, C (the case where one of them is ideal was dealt with above and anyway is not interesting). Let S be any circle through A and B, and let K be the midpoint of one of the arcs AB. Let KC meet S again in L. Then the perpendicular through L to KL meets AB in the required point D (Fig. 20).

For the angles ALK and KLB are equal, so that LK is the

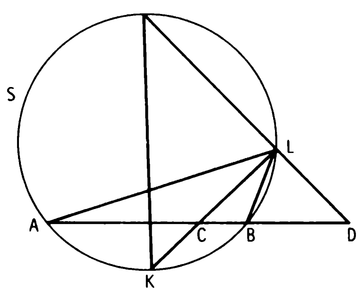

Fig. 20

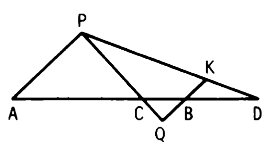

Fig. 21

interior bisector of the angle at L of the triangle ALB, and LD is the external bisector. See Example 3 above.

3. Draw parallel lines through A and B, and let them meet an arbitrary line through C (not parallel to them) in P and Q. On QB mark off $BK = QB$. Then PK meets AB in the required point D (Fig. 21). For K is the harmonic conjugate of Q with respect to B and the ideal point of the line QBK, and these four points are in perspective with A, B, C, D from the center P, so that the latter also form a harmonic set.

5.3. The Complete Quadrilateral and Quadrangle

A *complete quadrilateral* is defined to be the set of all those points of the projective plane that belong to one or other of four given lines, no three of which are concurrent. The lines are called the *sides* of the quadrilateral, and the six points of intersection of pairs of them are called the *vertices* of the quadrilateral. Two vertices are called *opposite* if they do not both lie on one side. Thus there are three pairs of opposite vertices; in an obvious terminology they are *ab* and *cd*, *bc* and *da*, and *ac* and *bd* (Fig. 22).

The three lines k, l, m joining pairs of opposite vertices are called the *diagonals* of the quadrilateral, and their three points of intersection are called the *diagonal points*. On each diagonal of a complete quadrilateral lie two vertices and two diagonal points.

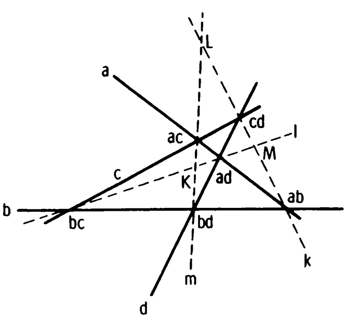

Fig. 22

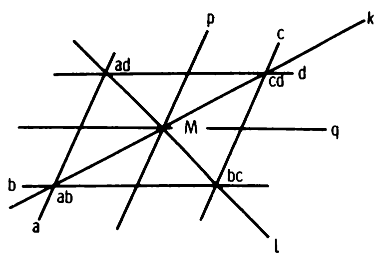

Fig. 23

Theorem 1. *Two opposite vertices of a complete quadrilateral, together with the diagonal points of the diagonal joining them, form a harmonic set.*

Proof. We will give a geometric proof using the first model of the projective plane. Suppose first that one of the diagonals, m for example, is the ideal line. Then the quadrilateral is a parallelogram, L is an ideal point, and M is the midpoint of the segment with end points ab and cd (Fig. 23). So the four points ab, cd, M, L form a harmonic set.

In exactly the same way, we see that the four points ad, bc, M, K form a harmonic set. Finally, the cross ratio of the four points ac, bd, K, L is, by definition, the cross ratio of the four lines p, q, l, k and so equals -1, by Example 8 of Section 5.1.

Suppose now that all three diagonals are ordinary. It may be proved (at least in the real projective plane) that the diagonals are not concurrent. So in this case they are not parallel, and therefore at least two of the diagonal points are ordinary. Suppose that K and L are ordinary diagonal points. Suppose that our projective plane Π is constructed from the Euclidean plane π. Consider another projective plane Π', constructed from a Euclidean plane π' that is parallel to k but not to π. Choose any point S of space in such a way that the plane through S and k is parallel to π'. Then the projection of Π onto Π' with center S takes $abcd$ into a quadrilateral $a'b'c'd'$ of Π' such that a' and b' are parallel to the line joining S to ab; c', d' are parallel to the line joining S to cd; and k' is the ideal line of Π'. Thus our image quadrilateral $a'b'c'd'$ is a parallelogram, and the theorem is already known in this case. Now cross ratio, and therefore harmonicity (in particular) is preserved by projection from S, so that the theorem is proved for the original quadrilateral. ▼

Let us define a *quadrangle* to be a set of four points of a projective plane, no three of them collinear. We also say the

points are the *vertices* of the quadrangle. The six lines joining the vertices two by two are called the *sides*, and two sides not through the same vertex are called *opposite*. In the complete quadrangle $ABCD$ there are three pairs of opposite sides—AB and CD, BC and DA, and AC and BD.

The three points of intersection of pairs of opposite sides are called *diagonal points* (Fig. 24), K, L, M, in the figure. It can be proved that in the real projective plane, they are not collinear. The lines joining the three diagonal points two by two are called the *diagonals*. Just as we may think of the diagonal points of a quadrilateral forming a *diagonal triangle*, so here we may think of the three diagonals forming a *diagonal trilateral*.

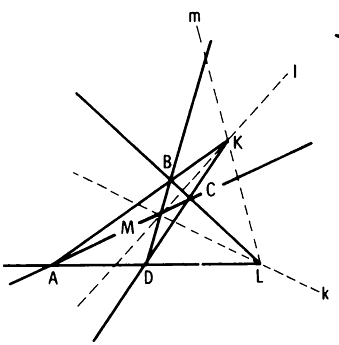

Fig. 24

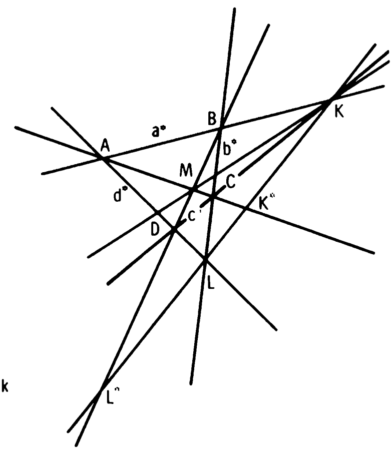

Fig. 25

Theorem 2. *Two opposite sides of a complete quadrangle together with the two diagonals through their diagonal point form a harmonic set.*

Proof. We will show, for example, that the four lines KA, KD, KM, and KL form a harmonic set (Fig. 25). Consider

the quadrilateral $a^*b^*c^*d^*$, where $a^* = AB$, $b^* = BC$, $c^* = CD$, and $d^* = DA$. For this quadrilateral, the points K and L are vertices, and if K^* and L^* are the diagonal points of $a^*b^*c^*d^*$ lying on KL, then $(KLL^*K^*) = -1$. Now (KLL^*K^*) is in perspective from B with $(ACMK^*)$, and this set is in perspective with KA, KC, KM, and KL. So they must all be harmonic sets, and the theorem is proved. ▼

Note. This theorem can be proved independently of the corresponding thereom for complete quadrilaterals, and, in fact, a proof analogous to the one we gave for quadrilaterals will work.

6. Examples of Projective Transformations

6.1. Hyperbolic Homology

Definition. Let α be a projective transformation of the projective plane Π, and suppose there is a line p of Π such that every point of it is fixed under α, and a point P, every line through which is its own image under α. Then α is called a *homology*. If P lies on p, α is a *parabolic homology*, and if not, α is a *hyperbolic homology*. Sometimes a hyperbolic homology is called an *elation*, and a homology (either kind) is called a *perspectivity*.

The line p is called the *axis* of the homology, and P is called the *center*. Note that in the definition of hyperbolic homology it is enough to define P as a point not on p invariant under α. For if l is any line through P, suppose that it meets p in K. Then the image of P is P, and that of K is K; since α preserves collinearity, any point on l must go into a point of $PK = l$, so that l is mapped into (of course, actually onto) itself.

Theorem 1. *There exists one and only one hyperbolic homology with given axis p and center P, such that the image of*

a given point A is a given point A' lying on AP (provided neither A nor A' is P or lies on p).

Proof. Suppose first that α is any such homology. We show that its action is uniquely determined. Let B be any point not on AA' or p (Fig. 26). Let AB meet p in M. Then $\alpha(M) = M$, so that A', B', and M are collinear, since A, B, and M are. Also, PB is its own image under α, so B' lies on PB. Then B' must be the point of intersection of $A'M$ and PB. Thus the action of α is determined on every point not lying on PAA'. But we can now repeat the argument, taking B and B' in place of A and A', to show that α is completely determined on every point not lying on PBB'. Since $\alpha(P) = P$, we have now completely determined α.

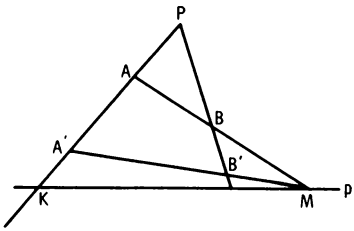

Fig. 26

The argument we have given so far is very general and works for any projective plane. The proof that α does, in fact, exist must be special, because there exist projective planes in which this is not true. For the real projective plane we use Theorem 1 of Section 3. Let C, D be any distinct points of p other than the point K of intersection of AA' with P. Then no three of the points P, A, C, D or P, A', C, D are collinear, so there exists a unique projective transformation β taking P, A, C, D into P, A', C, D. We wish to show that β fixes every point of p. But β fixes K, C, D, so, by the invariance of cross ratio, it must fix every point of p. It follows that $\beta = \alpha$. ▼

Theorem 2. *A hyperbolic homology may be represented as the product of a perspectivity and a rotation of the plane about a line. Conversely, the product of a perspectivity and a rotation is a hyperbolic homology.*

By this theorem we mean the following. Let π and π^* be planes in space that intersect on the line p. Then the transformation induced on Π by first taking a perspectivity of π onto π^* with center S and then rotating π^* into coincidence with π is a hyperbolic homology (with axis p). Conversely, given a hyperbolic homology α of Π, we may choose π^* through the axis p (regarded as a line of π rather than of Π), such that α is obtained by exactly the construction we have described (for a suitable choice of the center S of the perspectivity).

Suppose then (Fig. 27) that we define a transformation of Π by first taking the perspectivity γ of π onto π^* with center S, taking M into M^*, and then taking a rotation β of π^* about p into coincidence with π, taking M^* into M'. $\beta\gamma$ is certainly a projective transformation of Π, for γ is a projective mapping, and β is an affine mapping of π^* onto π that induces a projective mapping β of Π^* onto Π (Section 2 above), and the product of two projective mappings is projective.

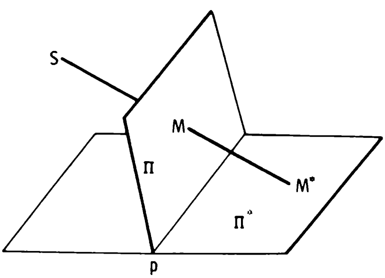

Fig. 27

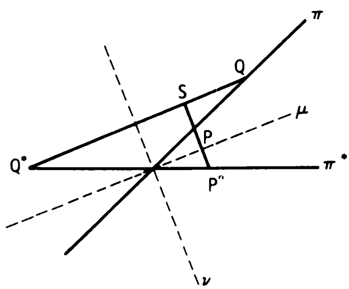

Fig. 28

We must show that there is a further fixed point, other than the points of the line p (which are all, including the ideal point, clearly fixed). To do this, we drop the perpendiculars from the center S of γ onto the angle bisector planes μ and ν of π and π^* (Fig. 28). If the perpendicular to μ meets π and π^* in P and P^*, and the perpendicular onto ν meets π and π^* in Q and Q^*, then clearly either P^* will go into P under β or Q^* will go

into Q. Since P goes into P^* and Q into Q^* under γ, either P or Q will be fixed under the composite mapping. The result now follows from Theorem 1.

Conversely, suppose that α is a projective transformation of Π that fixes P and every point of p. Let π^* be any plane (other than π) through p. For each point M of Π, let M' be its image, and let M^* be the image of M' in Π^* under the rotation β of π onto π^* about p. Let γ be the mapping of Π onto Π^* that takes each M into M^*. This mapping is projective, since $\gamma = \beta\alpha$, and α and β are both projective mappings.

We now have to show that γ is a perspectivity of Π onto Π^*. Let A, B, C be any three noncollinear points of Π, none of them lying on p, and let A', B', C' be their images under α. (Fig. 29). Then BC and $B'C'$ will meet in a point F of p, and, similarly, CA and $C'A'$ will meet in G on p, and AB and $A'B'$ in H on p. To prove this, suppose that F is the point in which BC meets p. Then F is fixed under α, so that the image $A'B'$

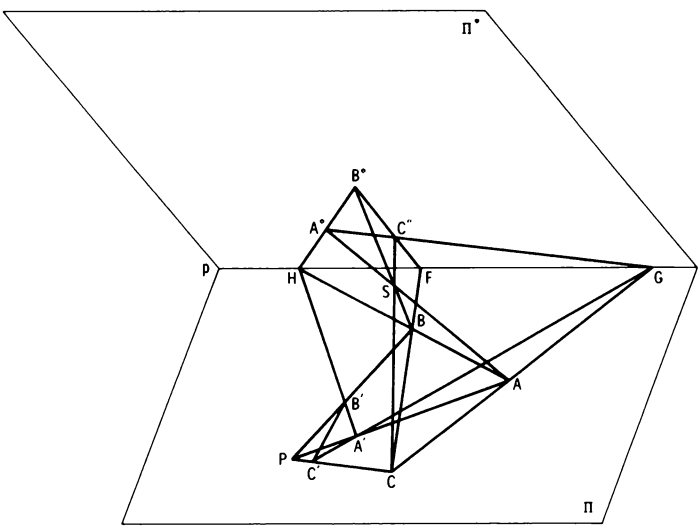

Fig. 29

of AB must pass through it. A similar result is obtained for the other points.

Suppose now that the rotation β takes A', B', and C' into A^*, B^*, and C^*, respectively. Then clearly the images B^*C^*, and so on, of the lines BC, and so on, will still pass through F, G, and H, respectively. So B, B^*, C, and C^* are all coplanar; similarly, C, C^*, A, and A^* are coplanar, and A, A^*, B, and B^* are coplanar. Since each pair of these planes has a line in common (for instance, BB^* is the common line of the planes BB^*CC^* and AA^*BB^*), no two of them are parallel, and therefore they intersect in a point S unless AA^*, BB^*, and CC^* are all parallel. Let us dismiss this case for the moment and let R be any point of p. Then the perspectivity with center S, like γ, takes A, B, C, and R into A^*, B^*, C^*, and R, respectively; since no three of these points are collinear, γ must be this perspectivity, and the theorem is proved.

If AA^*, BB^*, CC^* are parallel, then γ, like the parallel projection in the direction of AA^* of Π onto Π^*, would take A, B, C, R into A^*, B^*, C^*, R, so that they are equal, and γ is an affine transformation. But β is also an affine transformation, so $\alpha = \beta^{-1}\gamma$ is affine. Since α fixes every point of p and a further point P, it must be the identity. Now the identity can be regarded as a hyperbolic homology (in fact, if it is, the set of all hyperbolic homologies with given axis and center becomes a group), and in this case we may still make the theorem true by redefining π^* to be precisely π, and S to be any point not on it. ▼

Theorem 3. *A hyperbolic homology may also be represented as the product of two perspectivities.*

For suppose that α is a hyperbolic homology of Π with axis p and center P. Let π^* be any plane through p (other than π), and let l be any line through P and not lying in π. Let M and M' be any two points of π, not on p, such that M' is the image of M under α. Choose any point S_1 on l but not in π^*, and let S_1M meet Π^* in M^*. Suppose that $M'M^*$

meets l in S_2. Then α is the product of the perspectivities S_1 of Π onto Π^* followed by S_2 of Π^* back onto Π. For the product of these perspectivities is a projective transformation that leaves every point of p fixed and sends P into P and A into A'; so the result follows from Theorem 1. Note that the construction works, in the sense that $M'M^*$ does meet l. For P, M, and M' are collinear, so that the lines l and $MM'P$ are coplanar.

Conversely, if π and π^* are any two planes, and S_1 and S_2 are any two points not on either of them, then the perspectivity with center S of Π onto Π^* followed by the perspectivity with center S_2 of Π^* onto Π is a homology of Π. The axis of Π is the line of intersection of π and π^*, and its center is the point in which S_1S_2 meets Π. It is the identity if and only if S_1 and S_2 are the same point (in which case, the center is indeterminate). This result still holds if π and π^* are parallel (in which case, the axis is the ideal line of Π) or S_1S_2 is parallel to π (in which case, the center is an ideal point of Π). See Fig. 30.

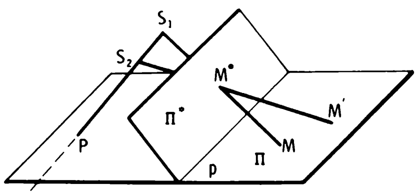

Fig. 30

6.2. Parabolic Homology

In the argument at the end of the last section, we did not consider the case in which S_1S_2 meets π on p. In this case there are clearly *no* fixed points of α except those of p. Nevertheless, our assertion that α is a homology with axis p and

center P is still true, for the reader may check that every line through P is its own image. Thus, in this case we have a parabolic homology (see beginning of Section 6.1). The situation is illustrated in Figs. 31 and 32.

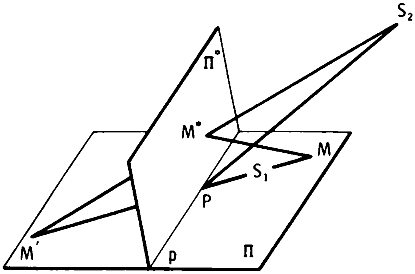

Fig. 31

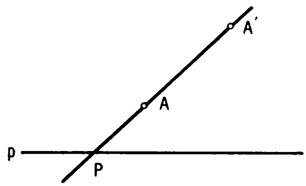

Fig. 32

Theorem 4. *If α is a projective transformation of Π under which every point of the line p is fixed, and there are no further fixed points, then α is a parabolic homology.*

Proof. We note first that if A is any point of Π not on p, and A' is its image under α, then AA' is fixed under α. For if AA' meets p in P, then $\alpha(AA') = \alpha(AP) = A'P = AA'$.

We now wish to show that every line through P is fixed under α. It is enough to show that the image B' of any point B lies on the line BP. For then every line BP through P will be fixed (as a whole) under α.

If B lies on AA', this is trivial, since AA' is its own image under α. If B does not lie on AA', the lines AA' and BB' are distinct. Since both are fixed under α, it follows that their point S of intersection is fixed by α (Fig. 33). It follows from our hypothesis that S lies on p. Thus $S = P$, and the theorem is proved. ▼

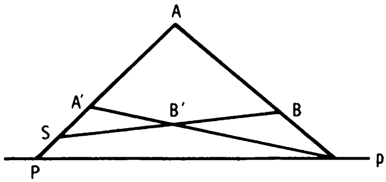

Fig. 33

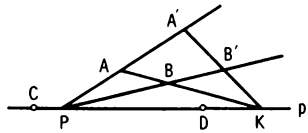

Fig. 34

Theorem 5. *There exists one and only one parabolic homology with given axis p and a given pair of distinct corresponding points A, A'.*

Proof. Let K be any point of p other than the point P of intersection of AA' with p. Let KA and KA' meet any line through P (other than p or AA') in B and B' (Fig. 34). Let C and D be distinct points of p, neither of them being K. Then it is clear that no three of the points A, B, C, D, and no three of the points A', B', C, D are collinear. So there exists a unique projective transformation β fixing C and D and taking A into A' and B into B'. Since β also fixes K, it fixes three points of p, and therefore every point of p, since it preserves cross ratio. So β is a homology. Its center must be the point of intersection of AA' and BB', P, which lies on p, so that β is parabolic. ▼

Theorem 6. *Any parabolic homology β of Π may be represented as the product of a perspectivity from Π onto Π^* (where π is any given plane through the axis of β) and a perspectivity from Π^* back onto Π.*

The proof is almost the same as for Theorem 3 above.

6.3. Hyperbolic and Parabolic Homologies in the Second Model

(A) *Hyperbolic Homologies.* We realize the projective plane as a bundle of lines and planes through a vertex S.

In this version, a hyperbolic homology turns out to be a transformation of the bundles such that:

1. Any three distinct lines lying in a plane are mapped into three distinct lines that also lie in a plane (condition for a projective transformation).
2. There is a fixed plane p of the bundle such that every line of the bundle that lies in p is mapped onto itself (p is the axis of the homology).
3. There exists a line P of the bundle, but not in p, which is its own image.
4. Every line is the image of some line (the mapping is onto).

Given any hyperbolic homology defined on a model of projective space of this kind, we may find an affine transformation of space leaving S fixed and generating this homology. This affine transformation is a skew compression onto p and in the direction of PS (Fig. 35). We allow negative coefficients: see Section 21.3 in Volume I.

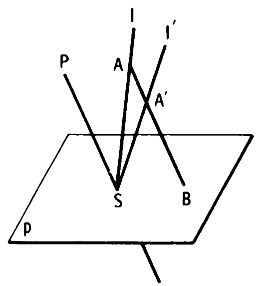

Fig. 35

Conversely, such an affine transformation generates a hyperbolic homology on the model.

Figure 26 may be regarded as a section of the bundle S made by a plane Π not through S. The point P (Fig. 26) is the point of intersection of the line PS (Fig. 35) and Π, and the line p (Fig. 26) is the line of intersection of the planes p (Fig. 35) and Π. Moreover, since P, SA, and SA' are coplanar, P, A, and A' are collinear (Fig. 26).

(B) *Parabolic Homologies.* We again interpret the projective plane as a bundle S. A parabolic homology turns out to be a transformation of the bundle such that:

1. The images of any three distinct lines of a plane are three distinct lines of some plane (condition for a projective transformation).

2. Every line of some plane p, and no other line, is mapped into itself (p is the axis of the homology).
3. The mapping is onto.

Given any parabolic homology defined on a bundle, there is an affine transformation of the plane leaving S invariant and inducing the homology; the affine transformation is a shear with "axis" p and in the direction of SP (the center of the homology). It is clear how we define a shear in space.

Conversely, any shear in space whose axis is the plane p through S and whose direction is that of PS induces a parabolic homology on the bundle S.

Figure 32 may be regarded as a section of the bundle S by a plane Π not passing through the point S (Fig. 36).

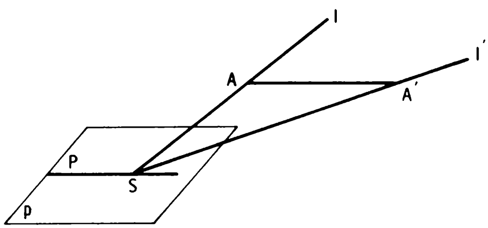

Fig. 36

6.4. Special Cases of Hyperbolic and Parabolic Homologies (First Model)

In this section we note some special cases of homologies, when the axis or the center, or both, are ideal.

1. If the axis of a hyperbolic homology α is the ideal line, then the restriction of α to π is an affine transformation, which is a homothetic transformation with center P, the center of the homology.

It is clear that α induces an affine transformation of π. Let A be any point of π other than P, and let A' be its image. Let B be arbitrary, and let B' be its image. Then P, B, B' are collinear, and P, A, A' are collinear, and AB is parallel to $A'B'$.

For in Π the lines AB and $A'B'$ must meet on the axis, which is the ideal line. It follows that

$$\frac{PB'}{PB} = \frac{PA'}{PA} = \text{constant.}$$

2. If the center of a hyperbolic homology α is an ideal point, then α induces an affine transformation of π that is a skew compression with axis p.

For, under a homology, any line through the center is mapped into itself. In particular, in this case, the ideal line is mapped into itself, so α induces an affine transformation of π. Since the transformation leaves invariant every point of p, it is an affinity (see Volume 1, Section 21.7). Since P is not the ideal point of p, the direction of the affinity is not parallel to its axis: that is, it is not a shear but a skew compression (see Volume 1, Section 21.7, Theorem 2).

3. A parabolic homology β whose axis is the ideal line induces an affine transformation on π. This transformation is a translation in the direction defined by the center of the homology.

For the axis is invariant under a homology, so that β maps π into itself and induces an affine transformation on π. If A and B are any two points, and A' and B' are their images, then AA' and BB' are parallel, since they both go through the center of the homology, which is an ideal point. Also AB and $A'B'$ are parallel, since their intersection in Π is a point on the axis. Thus $AA'BB'$ is a parallelogram, so that $\overrightarrow{AA'} = \overrightarrow{BB'} =$ a constant, and β is the translation through this vector.

4. If the center of a parabolic homology β is an ideal point, but the axis is not the ideal line, then β induces a shear on π.

We show as usual that β induces an affine transformation on π. If A and A' are corresponding points of π, AA' is parallel to p, since in Π it passes through the center P of the homology, and P is the ideal point of p. Moreover, as every point of p is invariant, the transformation must be a shear (See Volume 1, Section 21.7, Theorem 2).

6.5. Involutions of a Projective Plane

Suppose that we are given a point P of Π and a line p not passing through P. Given any point M of Π, we define a mapping α of Π into itself as follows. Let k be any fixed real number, or the symbol ∞. Let PM meet p in Q. Then we define $\alpha(M) = M'$ to be the unique point M' such that $(PQMM') = k$. The proof that this defines a unique point is quite simple. We introduce coordinates along the line PQ, and, supposing M' to have coordinate x, we write down the equation for x, which turns out to be linear. We define α to be the identity on P and p (where the other definition breaks down). Let us, for the moment, call α a *k-mapping*. Then we have:

Theorem 7. *Every k-mapping is a projective transformation, and, in fact, a hyperbolic homology. Conversely, every hyperbolic homology is a k-mapping for some k.*

Proof. Let γ be the hyperbolic homology that takes A into A' and has center P and axis p. If M is any point of the plane not on PAA', then its image M' under γ is given by the configuration of Fig. 37. Since $PRAA'$ and $PQMM'$ are in perspective from B, we have

$$k = (PRAA') = (PQMM').$$

Thus M' is the image of M under α. We can repeat the argument with M and M' playing the role of A and A' to show that α and γ coincide also on the points of AA'. Thus α and γ coincide.

Conversely, if γ is a hyperbolic homology and $\gamma(A) = A'$, set $k = (PRAA')$. We leave it to the reader to show that γ is the k-mapping with center P and axis p. ▼

Suppose now that $k = -1$. Then α^2 is the identity mapping (see Example 9, Section 5.1). Thus a hyperbolic homology that is a -1-mapping is a projective transformation,

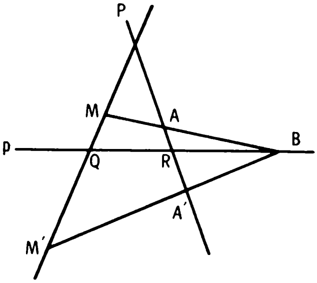

Fig. 37

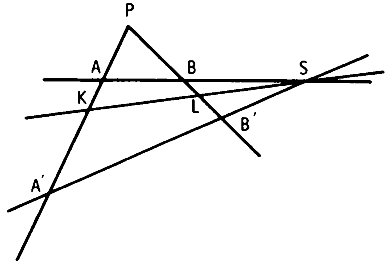

Fig. 38

not the identity, whose square is the identity. We call such mappings *projective involutions*. We can now state the converse of this result:

Theorem 8. *Any projective involution is a* -1*-mapping.*

Proof. Let α be a projective involution, not the identity. Then there are at least two points A, B whose images A', B' under α are distinct from them and from each other (Fig. 38). We leave the easy proof to the reader. Since α^2 is the identity, the image of A' is A and of B' is B. So the lines AA' and BB' are taken by α into themselves (since α is a projective transformation). Suppose that they meet in P. Then P is a fixed point for α. Let K be the harmonic conjugate of P with respect to A and A', and let L be the harmonic conjugate of P with respect to B and B'. Let K' and L' be the images of K and L under α.

Since projective transformations preserve cross ratio, $(AA'KP) = (A'AK'P) = -1$. But by Examples 9 and 10, Section 5.1, $(AA'KP) = -1$ says that K and P are harmonic conjugates with respect to A and A'. And $(A'AK'P) = -1$ says that K' and P are harmonic conjugates with respect to A and A'. Therefore, since the fourth harmonic to three given points is determined uniquely (Section 5.1), we conclude that $K' = K$, and, similarly, $L' = L$. So the image under α of the

line KL is the same line. Let S be the point of intersection of AB and KL. Then we show that $A'S$ passes through B'. Suppose that PB and $A'S$ (check that they are distinct lines!) meet in B''. Then $(AA'PK) = (BB''PL)$, both sets being in perspective from S. So $(BB''PL) = -1$, and since also $(BB'PL) = -1$, we must have $B'' = B'$.

S is an invariant point of KL. Since it is the point of intersection of AB and KL, its image is the point of intersection of $A'B'$ and KL, that is, S. Thus there are three fixed points K, L, S of KL. As α preserves cross ratio on KL, every point of KL must be fixed.

So α fixes P and KL, and is therefore a hyperbolic homology (see the remark above Theorem 1, Section 6.1). Since $(AA'PK) = -1$, it is a -1-mapping, by Theorem 7. ▼

7. Projective Transformations in Coordinates

7.1. The Fundamental Theorem

Theorem 1. *Suppose that we are given a coordinate system in the projective plane* Π *(see Sections* 1.2 *and* 1.4, *Vol.* 2). *Let* α *be a projective transformation of* Π *that takes the point* $M(x_1, x_2, x_3)$ *into the point* $M'(x_1', x_2', x_3')$. *Then the coordinates of* M' *are homogeneous linear functions of the coordinates of* M:

$$\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3; \\ x_2' &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3; \\ x_3' &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3; \end{aligned} \qquad (1)$$

where

$$\Delta = \Delta(\alpha) = \begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \neq 0. \qquad (2)$$

Proof. We first prove the converse theorem: The mapping defined by (1) and (2) is a projective transformation. Since (2)

holds, we may construct the inverse mapping, so that the mapping is one-one and onto. Next, the inverse image of a line

$$u_1'x_1' + u_2'x_2' + u_3'x_3' = 0$$

is also a line, given by

$$\begin{aligned} &u_1'(a_{11}x_1 + a_{12}x_2 + a_{13}x_3) \\ &\quad + u_2'(a_{21}x_1 + a_{22}x_2 + a_{23}x_3) \\ &\quad + u_3'(a_{31}x_1 + a_{32}x_2 + a_{33}x_3) = 0, \end{aligned}$$

or

$$\begin{aligned} &(u_1'a_{11} + u_2'a_{21} + u_3'a_{31})x_1 \\ &\quad + (u_1'a_{12} + u_2'a_{22} + u_3'a_{32})x_2 \\ &\quad + (u_1'a_{31} + u_2'a_{32} + u_3'a_{33})x_3 = 0. \end{aligned}$$

In this equation, the coefficients of x_1, x_2, x_3 are not all zero, because of (2). Thus the inverse image of a line is a line, and the transformation is projective. We show now that any projective transformation can be specified by a formula of the type of (1) above.

Suppose that α is a given projective transformation and that it takes $A(1, 0, 0)$, $B(0, 1, 0)$, $C(0, 0, 1)$ into the points $A'(a_1, a_2, a_3)$, $B'(b_1, b_2, b_3)$, $C'(c_1, c_2, c_3)$, respectively, (the triangle ABC is called the *triangle of reference*), and the point $D(1, 1, 1)$ into $D'(d_1, d_2, d_3)$ (D is called the *unit point*). We leave it to the reader to check that no three of A, B, C, D are collinear, so that the same holds of A', B', C', D' (since α is projective). It follows from the first fundamental theorem (Section 3 this volume) that α is completely determined by its effect on A, B, C, D. We wish to find a representation of α in coordinates. Consider the mapping β defined by

$$\begin{aligned} x_1' &= a_1\rho_1x_1 + b_1\rho_2x_2 + c_1\rho_3x_3\,; \\ x_2' &= a_2\rho_1x_1 + b_2\rho_2x_2 + c_2\rho_3x_3\,; \\ x_3' &= a_3\rho_1x_1 + b_3\rho_2x_2 + c_3\rho_3x_3\,; \end{aligned} \tag{3}$$

where none of ρ_1, ρ_2, ρ_3 are zero. This mapping, like α,

takes A, B, C into A', B', C', respectively. Let us try now to select the numbers ρ_1, ρ_2, ρ_3 so that β takes D into D'. For this to hold, we want

$$\begin{aligned} d_1 &= a_1\rho_1 + b_1\rho_2 + c_1\rho_3, \\ d_2 &= a_2\rho_1 + b_2\rho_2 + c_2\rho_3, \\ d_3 &= a_3\rho_1 + b_3\rho_2 + c_3\rho_3. \end{aligned} \tag{4}$$

Since A', B', C' are not collinear, we have

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \neq 0,$$

so that this set of equations has a nontrivial solution. We show that none of the ρ_i are zero. Suppose, for example, that $\rho_1 = 0$. Then

$$\begin{aligned} d_1 &= b_1\rho_2 + c_1\rho_3\,; \\ d_2 &= b_2\rho_2 + c_2\rho_3\,; \\ d_3 &= b_3\rho_2 + c_3\rho_3\,; \end{aligned}$$

which means that D' lies on the line joining B' and C'. This is false, by what we said above. Thus we have found ρ_i so that (3) gives a mapping β sending A, B, C, D into A', B', C', D'. Since no $\rho_i = 0$, the determinant of this mapping is nonzero:

$$\Delta(\beta) = \begin{vmatrix} a_1\rho_1 & b_1\rho_2 & c_1\rho_3 \\ a_2\rho_1 & b_2\rho_2 & c_2\rho_3 \\ a_3\rho_1 & b_3\rho_2 & c_3\rho_3 \end{vmatrix} = \rho_1\rho_2\rho_3\Delta \neq 0,$$

so that β is a projective transformation. By the first fundamental theorem, $\beta = \alpha$, and we have found that (3) is a representation for α. ▼

Theorem 2. *The representation of α in the form* (1) *is unique up to a constant multiple.*

By this we mean that if also

$$\begin{aligned} x_1'' &= a_{11}^* x_1 + a_{12}^* x_2 + a_{13}^* x_3; \\ x_2'' &= a_{21}^* x_1 + a_{22}^* x_2 + a_{23}^* x_3; \\ x_3'' &= a_{31}^* x_1 + a_{32}^* x_2 + a_{33}^* x_3 \end{aligned} \tag{5}$$

is a representation for α, then for some nonzero λ, $a_{ij}^* = \lambda a_{ij}$.

Proof. If (5) is a constant multiple of α and defines the mapping β, and $\alpha(M) = M'(x_1', x_2', x_3')$, then

$$\beta(M) = M''(x_1'', x_2'', x_3'') = M''(x_1'\lambda, x_2'\lambda, x_3'\lambda) = M'.$$

Thus $\beta = \alpha$, and any constant multiple of (1) is a representation for α.

Conversely, if (5) is *any* representation for α, we wish to show that the a_{ij}^* are constant multiples of the a_{ij}. Now, in the notation of Theorem 1, A' is the point $(a_{11}^*, a_{21}^*, a_{31}^*)$. Thus, for some ρ_1 we must have $a_{11}^* = \rho_1 a_1$, $a_{21}^* = \rho_1 a_2$, $a_{31}^* = \rho_1 a_3$. On considering the images of B and C similarly, we see that for some nonzero ρ_1, ρ_2, ρ_3 we must have (5) identical with (3). Since $\alpha(D) = D'$, furthermore, the ρ_i must satisfy the equation obtained from (4) when we write λd_i (for some nonzero λ) in place of d_i. For each value of λ, we obtain a unique solution, and, for given λ, the solutions are λ times the corresponding solutions of (4). On substituting back in (3) and comparing with (5), we see that $a_{ij}^* = \lambda a_{ij}$, as required. ▼

7.2. Invariant Points of a Projective Transformation

Theorem 3. *Any projective transformation α of Π has at least one fixed point P and one fixed line p (that is, a line p such that the image of any point of p is a point—perhaps a different point—of p).*

Proof. Suppose that α is given in coordinates by (1). The point $P(x_1, x_2, x_3)$ is fixed under α, if and only if $x_1' = \lambda x_1$,

$x_2' = \lambda x_2$, $x_3' = \lambda x_3$ for some $\lambda \neq 0$. Substituting these values in (1), we find

$$\lambda x_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3\,;$$
$$\lambda x_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3\,;$$
$$\lambda x_3 = a_{31}x_1 + a_{32}x_2 + a_{33}x_3\,;$$

or

$$\begin{aligned} (a_{11} - \lambda)x_1 + a_{12}x_2 + a_{13}x_3 &= 0\,; \\ a_{21}x_1 + (a_{22} - \lambda)x_2 + a_{23}x_3 &= 0\,; \\ a_{31}x_1 + a_{32}x_2 + (a_{33} - \lambda)x_3 &= 0. \end{aligned} \tag{6}$$

This system of equations has a nontrivial solution if and only if its determinant is zero:

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{vmatrix} = 0. \tag{7}$$

This is a cubic equation in λ with real coefficients, so that it has at least one real root, say λ_0. Since (2) holds, we cannot have $\lambda_0 = 0$. So, substituting λ_0 for λ in (6), we obtain a set of equations that determine the coordinates (x_1, x_2, x_3) of the corresponding P.

Suppose now that the line p is given by the equation

$$u_1'x_1' + u_2'x_2' + u_3'x_3' = 0.$$

The inverse image under α of this line is the line

$$\begin{aligned} u_1'(a_{11}x_1 &+ a_{12}x_2 + a_{13}x_3) \\ &+ u_2'(a_{21}x_1 + a_{22}x_2 + a_{23}x_3) \\ &+ u_3'(a_{31}x_1 + a_{32}x_2 + a_{33}x_3) = 0; \end{aligned}$$

or

$$\begin{aligned} (a_{11}u_1' &+ a_{21}u_2' + a_{31}u_3')x_1 \\ &+ (a_{12}u_1' + a_{22}u_2' + a_{32}u_3')x_2 \\ &+ (a_{13}u_1' + a_{23}u_2' + a_{33}u_3')x_3 = 0. \end{aligned}$$

In order that p should be its own image under α, therefore, it is necessary and sufficient that corresponding coefficients in these two equations should be proportionate, that is, that

$$a_{11}u_1' + a_{21}u_2' + a_{31}u_3' = \mu u_1';$$
$$a_{12}u_1' + a_{22}u_2' + a_{32}u_3' = \mu u_2';$$
$$a_{13}u_1' + a_{23}a_2' + a_{33}u_3' = \mu u_3';$$

for some $\mu \neq 0$, or

$$\begin{aligned}(a_{11} - \mu)u_1' + a_{21}u_2' + a_{31}u_3' &= 0;\\ a_{12}u_1' + (a_{22} - \mu)u_2' + a_{32}u_3' &= 0;\\ a_{13}u_1' + a_{23}u_2' + (a_{33} - \mu)u_3' &= 0.\end{aligned} \tag{8}$$

This system has a nontrivial solution if and only if its determinant is zero:

$$\begin{vmatrix} a_{11} - \mu & a_{21} & a_{31} \\ a_{12} & a_{22} - \mu & a_{32} \\ a_{13} & a_{23} & a_{33} - \mu \end{vmatrix} = 0. \tag{9}$$

This is a cubic equation in μ with real coefficients and therefore has at least one real root μ_0 and necessarily $\mu_0 \neq 0$. Substituting μ_0 for μ in (8), we obtain a system of equations that determine the coefficients u_1', u_2', u_3' of the equation of a fixed line. Note that since (9) and (7) coincide, they have the same roots.

This result is the best possible, in the sense that there exist projective transformations having only one fixed point and only one fixed line. Suppose, for example, that α is given by

$$\begin{aligned} x_1' &= \quad -x_2\,;\\ x_2' &= x_1\,;\\ x_3' &= \qquad\ \ x_3\,.\end{aligned}$$

We easily check that condition (2) is satisfied, so that α is a projective transformation. We may also check that it has the fixed point $C(0, 0, 1)$ and the fixed line c: $x_3 = 0$. We show first that there are no fixed points of c. If $P(a_1, a_2, 0)$ is a fixed

point of c, then it equals its image $P'(-a_2, a_1, 0)$. But this means that for some nonzero λ, we have

$$a_1 = -\lambda a_2, \qquad a_2 = \lambda a_1.$$

So $-\lambda^2 a_2 = a_2$. If $a_2 = 0$, then $a_1 = a_2 = 0$, which is impossible, since no point has coordinates (0, 0, 0). Thus $\lambda^2 = -1$, which is also impossible for a real number λ. So there are no fixed points on c. Suppose now that A is *any* fixed point other than C. Then CA is a fixed line and is not c (since C does not lie in c) and so meets c in a point X. But then X, the point of intersection of two fixed lines, is fixed, which we have shown to be impossible. Thus C is the only fixed point. We may prove similarly that c is the only fixed line, but we leave the proof as an exercise for the reader.

7.3. Homology Transformations in Coordinates

Let γ be a hyperbolic homology with center P, axis p, and a pair of corresponding points A and A'. We may choose a system of coordinates for Π such that P is the point (0, 0, 1) and p has equation $x_3 = 0$. Let us first consider for a moment what we are doing when we assign coordinates like this. Clearly, we are associating with the plane Π another projective plane in which coordinates already exist, for example, the fourth model (Section 1.6, this volume). By the first fundamental theorem (Section 3, this volume), there is a unique projective transformation of Π onto this plane Π', such that the images of four points, no three collinear, are the points A, B, C, D (defined as in the proof of Theorem 1 of Section 7.1 above) in Π'. We choose the four points to be two distinct points of p, P and a fourth point not collinear with any two of the others. Having established an isomorphism between Π and Π', we can, from now on, regard them as the same space (since for the purposes of projective geometry there is nothing to distinguish them) and thus may properly say that *P has coordinates* (0, 0, 1), rather than merely that it *is associated*

with the point (0, 0, 1) *of* Π'. If A is (a_1, a_2, a_3), and A' is (a_1', a_2', a_3'), then $(a_1', a_2', a_3') = (a_1, a_2, a_3')$, and γ is given in coordinates by

$$x_1' = x_1 \, ; \qquad x_2' = x_2 \, ; \qquad x_3' = kx_3 \, . \tag{10}$$

For the transformation given by (10) has the same effect as γ on $P(0, 0, 1)$ and every point $M(x_1, x_2, 0)$ of p, and it takes A into A' if we choose k in (10) to be $k = a_3'/a_3$ (check that neither of a_3 and a_3' is zero!). So α is the transformation given by (10), by the first fundamental theorem. In the particular case where $k = -1$, γ is an involution. In general, γ is a k-mapping, in the sense of Section 6.5, this volume. Conversely, any k-mapping is given by (10) in a suitable coordinate system, and any involution by (10) with $k = -1$.

Suppose now that β is a parabolic homology with center P, axis p, and corresponding points A, A'. We may choose coordinates so that P is (1, 0, 0), p is the line $x_3 = 0$, A is (0, 0, 1) (of course we take A not to lie on p), and A' is some point of the form $(a, 0, 1)$. We could even take A' to have coordinates (1, 0, 1), by taking the unit point to be some point on the line BA', where B is the point with coordinates (0, 1, 0). In any case, we see by the same argument as before that β is given in coordinates by

$$x_1' = x_1 + ax_3 \, ; \qquad x_2' = x_2 \, ; \qquad x_3' = x_3 \, . \tag{11}$$

Note. If we interpret the transformations γ and β given by (10) and (11) to refer to the second model of the projective plane, then γ is a skew compression onto the plane x_1Sx_2 in the direction of Sx_3 with coefficient k. In particular, if $k = -1$, it is the skew reflection in x_1Sx_2 in the direction of Sx_3. Similarly, β is a shear with respect to the "axis" plane x_2Sx_3 and parallel to the line Sx_1.

8. Quadratic Curves in the Projective Plane

A curve of the second degree, or a *quadratic curve*, of the projective plane is defined to be the totality of points whose

coordinates satisfy a homogeneous equation with real coefficients and of degree two in each of the coordinates:

$$a_{11}x_1^{\,2} + a_{22}x_2^{\,2} + a_{33}x_3^{\,2} + 2a_{23}x_2x_3 + 2a_{31}x_3x_1 + 2a_{12}x_1x_2 = 0. \quad (1)$$

We shall find it notationally convenient to write a_{32} for a_{23}, a_{13} for a_{31} and a_{32} for a_{23} upon occasion.

Notice that if one triple of coordinates for P satisfies this equation, so will any triple, and if one triple fails to do so, so will any triple fail. So our definition does not depend on the particular triple of coordinates chosen. If the equation had not been homogeneous, it would have been impossible to make such a definition, for some triples for P might satisfy the equation while others did not.

We now consider the interpretation to be put on quadratic curves in the first and second models. It will turn out that they are very closely similar to the ordinary conic sections.

8.1. Quadratic Curves in the First Model

Let (x_1, x_2, x_3) be a triple of homogeneous coordinates for an ordinary point of the projective plane (we are now thinking of Π as constructed from the ordinary plane π). Then the Cartesian coordinates for P are (x, y), where

$$x = x_1/x_3, \qquad y = x_2/x_3$$

(see Section 1.2 of this volume). It follows that the Cartesian coordinates of all the ordinary points of Π on the quadratic curve (1) satisfy

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0. \quad (2)$$

This is the equation of a quadratic curve of π, that is, a conic. The quadratic curves of Π given by (1) differ only in the additional ideal points they may have, that is, those points in which (1) cuts the line $x_3 = 0$. Note first that, conversely, if (2) is the equation of a conic of π, it consists of precisely the

ordinary points of the quadratic curve (1) of Π, so that any conic of π is part of a (unique) conic of Π. Let us now find the extra ideal points on C given by (1). The coordinates of such points must satisfy

$$x_3 = 0, \qquad a_{11}x^2 + 2a_{12}x_1x_2 + a_{22}x_2{}^2 = 0, \tag{3}$$

obtained by substituting $x_3 = 0$ in (1).

If

$$\delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}^2 > 0,$$

then (3) has only one real solution: $x_1 = x_2 = 0$, and this solution does not correspond to any point of Π, since there is no point with coordinates (0, 0, 0). Thus C consists only of ordinary points, and is, in fact, an ellipse (or a single point or nothing). This satisfies our intuitive demand that the ideal points of C should be in some sense the "infinitely distant" points on (2); in the case of an ellipse, which is contained in a bounded portion of the plane, there are no "infinitely distant" points.

If $\delta < 0$, Eq. (3) has two distinct real roots, so that the ideal line $x_3 = 0$ meets C in two distinct ideal points. In this case, (2) is a hyperbola or two intersecting lines, and we may think of (1) as representing this hyperbola together with the two "infinitely distant" points on it, that is, the ideal points of its two asymptotes. If (2) is a pair of intersecting lines, we may think of it as a special case of the hyperbola, in which it has become its asymptotes, and the two ideal points of C are the ideal points of these two lines.

Suppose finally that $\delta = 0$. Then (3) has a unique nonzero solution $x_1 : x_2 : x_3 = -a_{12} : a_{11} : 0 = a_{22} : -a_{12} : 0$, and the ideal line $x_3 = 0$ meets (1) in a single ideal point. In this case, (2) represents a parabola or the degenerate forms of a parabola—two parallel lines, or two coincident lines. (1) is then the parabola, together with the ideal point on its diameter. We may think of it as an "infinitely elongated" ellipse, touching the ideal line at the ideal point on its diameter, which is the

vertex opposite the (ordinary) vertex. If it is two parallel or two identical lines, the ideal point of (1) is the ideal point of these lines.

We see that the classification of affine conics can be made while keeping one eye on their projective analogs. Thus an ellipse is a conic with center at an ordinary point, not meeting the ideal line; a hyperbola is a conic with center at an ordinary point, intersecting the ideal line; a parabola has center on the ideal line, touching it. In the degenerate cases, we interpret the center of a pair of intersecting lines to be their point of intersection, whether they are real or imaginary lines.

So far, we have considered only those cases where at least one of a_{11}, a_{12}, a_{22} is nonzero, so that (2) is genuinely the equation of a conic and not just of a line. In the special case where all three are zero, Eq. (1) becomes

$$2a_{13}x_1x_3 + 2a_{23}x_2x_3 + a_{33}x_3{}^2 = 0.$$

This represents a conic consisting of two lines: $x_3 = 0$ and the line $2a_{13}x_1 + 2a_{23}x_2 + a_{33}x_3 = 0$. In the special case where $a_{13} = a_{23} = 0$, (1) represents the ideal line "taken twice over."

It is not hard to show now the following elegant result:

Theorem 1. *The quadratic curve* (1) *is degenerate if and only if the associated determinant*

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

is zero.

A quadratic curve is said to be *degenerate* if it splits into two lines, or, algebraically speaking, if the expression (1) for it factors. We allow complex coefficients in the factorization, so that we must consider a single point (a pair of imaginary intersecting lines) as degenerate, whereas the empty set (an imaginary ellipse) is not.

8.2. Quadratic Curves in the Second Model

Let x_1, x_2, x_3 be the projective coordinates of a "point" in the second model. Then it is known, from analytic geometry of three dimensions, that Eq. (1) defines a cone through the origin S; that is, it consists entirely of rays through S and is such that the intersection with any plane not through S is a conic. We distinguish five cases:

1. The cone is nondegenerate; that is, at least one plane intersects it in an oval conic.
2. It consists of a pair of distinct planes.
3. It consists of a pair of coincident planes.
4. It is an "imaginary nondegenerate cone"; that is, it consists of just the origin.
5. It is a "pair of imaginary planes"; that is, it consists of a single line through S.

We may get back to the first model by considering the intersection of this second model with any plane π not through S. Clearly, any two such planes π, π' are in perspective from S, and the conics in which they, or rather the corresponding Π and Π', meet a given cone C of S are projectively equivalent. We consider cases:

1. We may choose π to meet C in an ellipse, a parabola, or a hyperbola. It follows that these three types of conics (the oval conics) are projectively equivalent.
2. We may choose π to meet C in distinct lines or parallel lines (the latter, if π is parallel to the line of intersection of the two planes). Thus distinct and parallel lines are projectively equivalent.
3. Any π meets C in a line ("two coincident lines").
4. Any π meets C in the empty set.
5. Any π, or rather Π, meets C in a single point.

We thus see that there are just five projective classes of conics (compare Volume 1, Section 31). We note that the distinction between the three kinds (affine classes) of oval curves and also the distinction between intersecting and parallel

lines have disappeared. This is reasonable, for in the first model we can only distinguish them by reference to the ideal line; an oval curve is an ellipse, parabola, or hyperbola according to whether it meets the ideal line in no point, one point, or two points. And we distinguish parallel from intersecting lines according to whether they meet on the ideal line or not. In "pure" projective geometry, where there are no ideal points or lines, these distinctions disappear. We examine the situation in greater detail in the next part.

8.3. Quadratic Curves in the Fourth Model

We now give a projective classification of the projective quadratic curves, by considering them in the fourth model. By "projective classification" we mean a classification into types such that any two curves of the same type are projectively equivalent (that is, may be obtained from each other by projective transformations), and no two curves of different types are projectively equivalent. Compare Volume 1, Section 31, where we did the same for affine quadratic curves.

It is known, from algebra, that any quadratic form

$$a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{23}x_2x_3 + 2a_{31}x_3x_1 + 2a_{12}x_1x_2 \tag{4}$$

can be reduced by a nonsingular linear transformation

$$\begin{aligned} x_1 &= c_{11}X_1 + c_{12}X_2 + c_{13}X_3\,; \\ x_2 &= c_{21}X_1 + c_{22}X_2 + c_{23}X_3\,; \\ x_3 &= c_{31}X_1 + c_{32}X_2 + c_{33}X_3 \end{aligned} \tag{5}$$

into the form

$$\varepsilon_1 X_1^2 + \varepsilon_2 X_2^2 + \varepsilon_3 X_3^2, \tag{6}$$

where ε_1, ε_2, ε_3 assume values -1, 0, or $+1$. According to a famous theorem known as Sylvester's law of inertia or the invariance of signature, whatever transformation (5) we choose, the number of $+1$'s, -1's, and zeros in (6) will be the same. Geometrically, this theorem states that any curve (1) can be reduced by a projective transformation to one and only one of the following five forms (see Theorem 1, Section 7, this volume):

(1) $X_1{}^2 + X_2{}^2 - X_3{}^2 = 0$ (oval curve);
(2) $X_1{}^2 - X_2{}^2 = 0$ (two distinct lines);
(3) $X_1{}^2 = 0$ (two coincident lines);
(4) $X_1{}^2 + X_2{}^2 + X_3{}^2 = 0$ (imaginary oval curve);
(5) $X_1{}^2 + X_2{}^2 = 0$ (two imaginary lines).

This is just what we would expect from our examination of the second model, and also of the first when we refuse to give the ideal line any special status.

8.4. Internal and External Points of a Projective Conic

Definition. The point M in Π is said to be an *internal* (or *interior*) point of the oval conic K (case 1, Section 8.3 above), provided that every line through M meets K in two distinct points; a point N is said to be *external* to K (or an *exterior* point of K) provided there exist lines through N not meeting K at all.

Theorem 2. *A given point M of Π has exactly one of the following properties:* (a) *it lies on K;* (b) *it is an internal point of K;* (c) *it is an external point of K.*

Here K is a given oval conic.

Proof. We can find a plane π' and a circle C in π' such that Π is the image of Π' under some projective transformation α and K is the image under α of C. Let M be any point of Π, and suppose its original in Π' is M'. If M' lies on C, then the tangent t' to C through M' meets C in only one point—M' itself. But then the image of t' under α is a line t passing through M and through no other point of K. So M is a point of K (since M' is a point of the original C of K) and is not an external or an internal point (for *every* line through M meets K in one point, M, and there is a line t which does not meet it in any other). We call t the *tangent* to K at M.

Suppose next that M' is inside C (in the usual sense of the

word). Then every line through M' meets C in two distinct points, and the same must be true for the lines through M. So M is an internal point; it is not external (since it is impossible from the definition for a point to be both internal and external), and it is not a point of K (since M' is not a point of C). Suppose finally that M' is outside C, that is, either outside in the usual sense or an ideal point. Then there is a line l' through M' that does not cut C, for example, the perpendicular to the radius through M', if M is an ordinary point, and the ideal line otherwise. Then l is a line through M not meeting K, so that M is external, and it is not internal (by definition) and not on K (since M' is not on C). The theorem is thus completely proved. ▼

Theorem 3. *Projective mappings preserve interiors and exteriors of oval conics.*

This means the following. Let K be an oval conic of the projective plane Π, and let K' be an oval conic of Π'. Let α be a projective transformation of Π onto Π' taking K into K'. Then the image under α of an internal point of K is an internal point of K', and the image of an external point of K is an external point of K'. Since α is onto Π', we have, as a corollary, that if α maps K into K', then, in fact, it maps K *onto* K'. For every point of K' is the image of some point of Π, and since this point cannot be an external or an internal point of K, it must be a point of K itself, by Theorem 2.

The proof follows closely the lines of the proof of Theorem 2, and we leave it to the reader. Notice that we have also proved the following:

Definition. The line l is said to be a *tangent* at the point M of an oval conic K, provided that l is a line through M meeting K in M and no other point.

Theorem 4. *There is one and only one tangent at each point of an oval conic.*

Proof. There is one and only one tangent (in this sense) at each point of an affine circle. We leave further details to the reader; they follow the proof of Theorem 2 closely. ▼

9. Projective Transformations of Space

9.1. PROJECTIVE 3-SPACE

We saw in this chapter how to turn the ordinary plane into a projective plane, which has many nice properties, by adjoining extra points (which we called "ideal") and regarding all the ideal points as belonging to an extra line, the ideal line.

In exactly the same way, we shall now adjoin extra "ideal" points, extra "ideal" lines (consisting of special points), and an extra "ideal" plane (consisting of all the special points) to ordinary Euclidean space, in order to obtain an object called *projective space*, or, more specifically, *real projective 3-space* (*real* because in the coordinate version, model 4, we shall have coordinates that are real numbers, and *3* because it is constructed from three-dimensional space).

Specifically, to each sheaf of parallel lines in space S we assign one extra "point," which we regard as lying on all the lines of the sheaf. We assign different "points" to distinct sheaves. We shall call all these extra "points" the "ideal points". If π is a plane of S, the set of all the ideal points lying on lines of π will be regarded as an ideal line. We leave it to the reader to show that parallel planes have the same ideal line. Finally, we regard all the ideal points as forming an ideal plane.

We now define the objects we shall have in our projective space. First are the (projective) *points*, consisting of all the ordinary points and all the ideal points. Next are the *projective lines*. A set of points will be called a projective line if it consists of all the points of an ordinary line, together with the ideal point of that line (that is, attached to the sheaf of parallel lines of which that line is a member), or consists of all the ideal points of an ordinary plane (as in the definition we gave above).

Finally there are the *projective planes.* A set of points will be regarded as a projective plane if it consists of all the points of an ordinary plane, together with the ideal line of that plane, or if it consists of all the ideal points (and no others). We call this latter the "ideal plane." The collection of all the points, ordinary and ideal, together with this structure of lines and planes, is called *projective space.*

We now say what we mean by "lying on" and "passing through." We say a point *lies on* a line or a plane if it is a member of the line or the plane (regarded as a set of points). We say a line or a plane *passes through* a point if the point is a member of it (that is, lies on it). We say a plane passes through a line if the line is a subset of it. We can contract this notation by using the word "incident." Thus we say that a point and a line, a line and a plane, or a point and a plane are *incident* if the first lies on the second. We shall denote points by capital letters A, B, ..., lines by italic letters l, m, ..., and planes by Greek letters α, β. We shall write AB for the line through A and B, "the line incident to A and B," and $\alpha\beta$ for the line of intersection of the planes α and β, "the line incident to α and β." We now list the "nice" properties of projective space that inspired our construction:

- **I.** *There is one and only one line incident to two given distinct points.*
- **II.** *There is one and only one line incident to two given distinct planes.*
- **III.** *If the distinct points A, B, are incident to a given plane, so is the line AB.*
- **IV.** *If the distinct planes α, β, are incident to a given point, then so is the line $\alpha\beta$.*
- **V.** *There is at least one plane incident to all of three given points.*
- **VI.** *There is at least one point incident to all of three given planes.*

In order that the reader should gain some insight into the reason why we constructed projective space as we did, we

recommend very strongly that he give detailed proofs of all these properties.

In exactly the same way as we defined an abstract projective plane (Section 1.6, this volume), we can use the properties I to VI above to define an abstract *projective 3-space*. That is, we define a projective 3-space to be a collection of objects called points, of certain nonempty sets of points called lines, and of nonempty sets of points called planes, such that I to VI are satisfied (where "incident" means what we explained it to mean earlier). We must add an extra axiom:

L. *There exist five points, no four of which are coplanar.*

(that is, incident to the same plane). In such an abstract space, there are no ideal points or lines; all are on the same footing. What the reader will have shown in proving the properties I to VI of the space we constructed from Euclidean 3-space is that this construct is, in fact, an example of a projective 3-space. We can give other models of this space, which correspond to the second and third model of the projective plane, but they require us to consider sheaves or spheres in four-dimensional Euclidean space, which the reader may consider more complicated than is worthwhile. However, in Section 9.3 below, we give the fourth (coordinate) model of real projective space, and it is no more complicated than the fourth model for the plane. There exist projective 3-spaces not isomorphic to the real projective 3-space we have constructed; for instance, *complex* projective 3-space, obtained by using complex instead of real coordinates in the fourth model.

Using only the axioms I through VI and L for an abstract projective space, it is possible to prove a number of other similar results in any projective space. We list five such:

L′. *There exist five planes, no four of which are incident to a single point.*

VII. *Any two distinct lines incident to the same plane are incident to one and only one point.*

VIII. *Any two distinct lines incident to the same point are incident to one and only one plane.*

IX. *A point and a line not incident to it are together incident to one and only one plane.*

X. *A plane and a line not incident to it are together incident to one and only one point.*

We now prove some of these results.

Proof of L′. Let 1, 2, 3, 4, 5 be five points no four of which are coplanar. By V we may find planes through 123, 124, 145, 235, 345, respectively. Clearly none of these planes contains a fourth one of the five points, so that all the planes are distinct. Suppose we are given a set of four of these planes, say α, β, γ, δ. It will be found that it is possible to select three of them, say α, β, γ, having some point A in common in such a way that two of these, say α, β, have a second point B in common (where A and B are each one of $1, \ldots, 5$).

By III α and β have the line AB in common, and by II this is the *only* line they have in common, since they are distinct planes. Thus if there is a line common to α, β, γ, it must be AB. But then B would lie on γ, which is false. So by III α, β, γ have only the point A in common. Since A does not lie on δ, our four planes α, β, γ, δ have no point in common. Thus the five planes we have listed satisfy the requirement of L'. ▼

Lemma. *Every line contains at least two points.*

Proof. Suppose on the contrary the line m contains the single point M. By L Σ contains at least two points besides M, say A and B. By V there is a plane π through M, A, B. By L again, there exists some point C not on π: let π' be a plane through M, A, C. By III the line $a = MA$ is incident to both π and π'; since these planes are distinct, this is the *only* line incident to both of them (by II). But the line m is incident to both of them. Thus $m = a$ contains a second point A besides M. ▼

Proof of IX. Let m be a line and M a point not on it. By the lemma we may choose distinct points A, B, on m. Then

by V there is a plane π incident to all of A, B, M. By III m is incident to π.

We leave the uniqueness part to the reader (use II and III). ▼

At this point the reader may easily prove VIII.

Proof of VII. Let m and n be distinct lines both incident to the plane π. By L choose a point A not on π. Then by IX we may find a plane π_1 incident to m and A, and a plane π_2 incident to n and A. By VI the three planes have a common point P. If, for example, P is not incident to m, then the distinct planes π and π_1 are both incident to m and P, contradicting the uniqueness part of IX. So P is incident to m, and similarly to n.

Uniqueness follows immediately from I. ▼

9.2. The Principle of Duality

Let us pair the properties I to X as follows: I with II, III with IV, V with VI, L with L′, VII with VIII and IX with X. We see that each proposition of a pair is obtained from the other by interchanging the words "point" and "plane" and keeping everything else as it is (including the words "line" and "incident"). It was to get this tidy correspondence that we introduced the word "incident"; otherwise we should have had to interchange "lie on" and "pass through" as well.

Two propositions about points, lines, and planes, formulated entirely in terms of incidence, are called *dual* if they may be obtained from each other by interchanging "point" and "plane," leaving everything else alone. We thus see that our ten propositions about a projective plane come in dual pairs.

Suppose now that we have a proof of some theorem that holds in any projective space; for example, our proof of VII above. If we go through the proof, systematically replacing every statement by its dual, we obtain a proof of the dual statement (in our case VIII). For at each stage of the original proof we are either using a logical argument, or appealing to one of the axioms. So at the corresponding stage of the dual of the proof we are either again using a logical argument, or appealing to the dual of an axiom. Now the dual of an axiom is

true (either it is itself an axiom or it is L′, which we have proved). It follows that the dual of a proof uses correct arguments from true premises, and so leads to true conclusions. Thus the dual of any theorem for an abstract projective space is again a theorem. This principle is known as the *principle of duality*; for a more detailed discussion of it, see Appendix 2, p. 126.

We may speak not only of dual statements but also of dual figures. Thus we have seen that in the projective plane (where dual statements are obtained by interchanging "point" and "line"), quadrangles and quadrilaterals are dual, and a triangle defined as three noncollinear points (the vertices) together with the lines through pairs of them (the sides), is self-dual, as the reader may easily check. Two figures are dual if they are defined by conditions on the existence and number of points, lines, and planes comprising the figure, together with various incidence relations, and these conditions are dual statements.

Consider, for example, the cube, thought of as a configuration in projective 3-space. We may think of it roughly as a set of eight points lying four by four (in a certain arrangement) on six planes and joined pairwise by twelve edges in a certain way. The dual figure is an octahedron (Fig. 39). We may see this by arguing as follows: A cube has six faces, so the dual will have six vertices. Each vertex of the cube has three edges incident with it, so each face of the dual will have three edges incident with it, that is, each face will be a triangle. Next, each face of the cube has four vertices, so that each vertex of the dual will be incident to four faces. With such information we can construct the octahedron. So far, we have not considered the regularities of the cube that distinguish it from other figures with the same arrangement of faces, edges, and vertices—we might so far be dealing with a very squashed box, and our octahedron might be very squashed too. We can introduce some regularity by considering, for example, that opposite edges of the cube are coplanar, so that opposite edges of the octahedron are concurrent (in the figure, they meet in an ideal point). Or we could notice that the joins of opposite vertices in the cube are concurrent, so that the intersections of opposite

faces of the dual are concurrent. In our figure, these lines are all ideal lines, concurrent in an ideal point. The reader may be able to find more "dual" regularities of the regular cube and the regular octahedron.

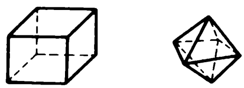

Fig. 39

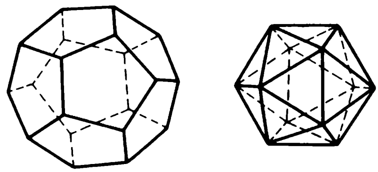

Fig. 40

We can consider another of the regular solids, the dodecahedron, which has twelve faces, each a regular pentagon, thirty edges, and twenty vertices. The dual figure will have twelve vertices, thirty edges and twenty faces—such a figure is the regular icosahedron (Fig. 40). In the dodecahedron three edges are concurrent at each vertex, and correspondingly in the dodecahedron there are three edges on each face (that is, each face is a triangle). In the icosahedron, each face has five edges, and correspondingly in the icosahedron there are five edges through each vertex. In both the regular icosahedron and the regular dodecahedron the diagonals are concurrent, so in each of them the lines of intersection of opposite faces are coplanar (the plane is the special plane in our figure).

Let us note finally that the tetrahedron is self-dual.

9.3. Homogeneous Coordinates

We will first show how to introduce homogeneous coordinates in the first model of real projective space S. Introduce Cartesian coordinates in space, and suppose that the point M has coordinates (x, y, z). Form the quadruple $(x, y, z, 1)$, and consider all the quadruples (x_1, x_2, x_3, x_4) proportional to it:

$$x_1 : x_2 : x_3 : x_4 = x : y : z : 1;$$

that is,

$$x_1 = \lambda x, \qquad x_2 = \lambda y, \qquad x_3 = \lambda z, \qquad x_4 = \lambda,$$

where λ takes all real values except zero. For any such quadruple, since $x_4 \neq 0$, we have

$$x = x_1/x_4, \qquad y = x_2/x_4, \qquad z = x_3/x_4.$$

Let us assign to M, thought of as a point of projective space, *all* these proportionate quadruples. We call any one such quadruple *homogeneous coordinates* for M. If M is an ideal point of S, let **a** be a vector parallel to every line of the bundle associated with M, and suppose that its coordinates are X, Y, Z. Then we assign as homogeneous coordinates to M all those quadruples (x_1, x_2, x_3, x_4) such that

$$x_1 : x_2 : x_3 : x_4 = X : Y : Z : 0,$$

that is,

$$x_1 = \lambda X, \qquad x_2 = \lambda Y, \qquad x_3 = \lambda Z, \qquad x_4 = 0$$

We may check that this definition is independent of the particular choice of the vector **a**.

We may now check that any projective plane is given by a homogeneous linear equation,

$$u_1 x_1 + u_2 x_2 + u_3 x_3 + u_4 x_4 = 0; \tag{1}$$

a point lies on this plane if and only if any quadruple of homogeneous coordinates for it satisfies this equation. Conversely, any such linear equation defines a projective plane. In particular, the ideal plane has equation $x_4 = 0$.

Having assigned coordinates to the first model, we can now, just as in Section 1.6 of this volume, see how to define the fourth model.

In the fourth model for projective space we take the points to *be* all the complete sets of proportionate quadruples. If M is a point and (x_1, x_2, x_3, x_4) is a quadruple of M, we will speak loosely of (x_1, x_2, x_3, x_4) as being *homogeneous coordinates for M*. We define a plane to be the set of all those points satisfying some equation (1), where not all the u_i are zero. We may check that this is an unambiguous definition, in the sense that, for a given point M, either every quadruple or no quadruple of homogeneous coordinates satisfies this equa-

tion. We define a *line* to be the set of points of intersection of two distinct planes. It is then a matter of direct algebra to verify that all the properties I through X hold, so that we have indeed constructed a projective space. The assignment of coordinates to the first model, which we made at the beginning of this section, was really a proof that the first and fourth models are isomorphic. The reader is recommended to carry out this proof in detail. The fourth model is "pure," in the sense that there are no longer any ideal points or lines. Of course, in the correspondence we set up between the first and fourth models we associated the ideal plane with the plane $x_4 = 0$, but this is accidental; we might easily have chosen some other association. Using this model, we can now show that real projective space R is self-dual. We note† that the dual space R^* of R is obtained by taking the points of R^* to be the planes of R, and the planes of R^* to be the points of R, with the definition that a point α of R^* is incident with the plane P of R^* if and only if the point P of R is incident with the plane α of R. The lines of R^* are the same as the lines of R, but are thought of as the set of planes (of R) through the line rather than as the set of points (of R) on them.

Let us define a mapping of R^* onto the fourth model of R. If α is a point of R^*, it has Eq. (1), regarded as a plane of R. We assign it the coordinates (u_1, u_2, u_3, u_4). Since (1) can be multiplied throughout by any nonzero λ without changing the set of points satisfying the equation, α also has any proportionate quadruple $U = (\lambda u_1, \lambda u_2, \lambda u_3, \lambda u_4)$ as a set of coordinates. Conversely, if V is not proportional to U, the equation corresponding to V as (1) does to U does not represent the same plane, so that V under our mapping will be the coordinates of a different point β of R^*. It is clear that any nonzero quadruple U defines an equation (1), and this is the coordinates of some point α of R^*. So the assignment of coordinates that we have defined gives a one-one mapping of the points of R^* onto the points in the fourth model of R. We have to show that this mapping preserves collinearity. Let U, V, W be three

† See Appendix 2, p. 126.

collinear points of R^*. This means that the corresponding Eqs. (1), and so on, in R represent three planes through a line. Now the condition for this is that there exist real numbers λ, μ, not both zero, such that $w_i = \lambda u_i + \mu v_i$ $(i = 1, 2, 3, 4)$. But this condition is precisely the same as that in the fourth model of R for the three points (u_1, u_2, u_3, u_4), and so on, to be collinear. We have thus established a projective isomorphism between R^* and the fourth model of R, by a very natural method of assigning coordinates. Note that Eq. (1), interpreted in R^* with the u_i variable and the x_i fixed, gives the equation of the plane M of R^*. Here M is the point (x_1, x_2, x_3, x_4) of R, and (u_1, u_2, u_3, u_4) are the coordinates (in R^*) of all the planes of R incident with M.

A very similar, but simpler argument, establishes that the projective plane is self-dual. We leave the formulation and proof of this to the reader.

9.4. Projective Transformations of Space

We define a projective transformation of projective space to be a one-one mapping of the space onto itself, under which the images of any three collinear points are collinear, The set of projective transformations of space is a group. Moreover, the image of a plane is a plane, and the mapping restricted to this plane is a projective mapping of one projective plane onto another.

Theorem I. *Given any five points A', B', C', D', E' of projective space, no four coplanar, and five points A'', B'', C'', D'', E'', no four coplanar, there exists one and only one projective transformation of space taking A' into A''*, and so on.

The proof of the corresponding theorem in the plane (Section 3.1, this volume) cannot be extended to this case, and we give a different proof in Section 9.5.

As in the projective plane, cross ratio is preserved by projective transformations of space. This follows at once from the remark above that a projective transformation of space induces a projective mapping of every plane.

9.5. Projective Transformations in Coordinates

Let R be the fourth model of projective space, and let α be a projective transformation. Suppose that for each $M = M(x_1, x_2, x_3, x_4)$ we have $\alpha(M) = M' = M'(x_1', x_2', x_3', x_4')$.

Theorem 2. *α may be specified in coordinates by the equations*

$$\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4, \\ x_2' &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4, \\ x_3' &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{24}x_4, \\ x_4' &= a_{41}x_1 + a_{42}x_2 + a_{43}x_3 + a_{44}x_4, \end{aligned} \tag{1}$$

subject to the condition that

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} \neq 0. \tag{2}$$

The representation is unique up to a scalar factor (that is, every a_{ij} may be multiplied by the same factor, without altering α, but otherwise the representation is unique). *Conversely, any set of Eqs.* (1), *subject to* (2), *defines a projective transformation of space.*

The proofs are almost identical to those of Section 7.1 of this volume.

We can use this representation to prove Theorem 1. Of course, if we follow the proof of Theorem 1, Section 7.1, we will find ourselves assuming what we are supposed to prove. So we proceed a little differently.

Proof of Theorem 1. Let $A(1, 0, 0, 0)$, $B(0, 1, 0, 0)$, $C(0, 0, 1, 0)$, $D(0, 0, 0, 1)$ be the vertices of the tetrahedron of reference, and let $E(1, 1, 1, 1)$ be the unit point. Let A', B', C', D', E' be five points, no four of them collinear. We will try to construct a mapping, given by (1), that takes A, B, C, D, E into A', B', C', D', E', respectively. Suppose that such a mapping exists. Then $a_{11}, a_{21}, a_{31}, a_{41}$ are determined up to a scalar multiple ρ_1 as the coordinates of A'—and similarly for the other

columns. To force $\alpha(E) = E'$, we have a system of equations for $\rho_1, \rho_2, \rho_3, \rho_4$ parallel to Eqs. (4) of Section 7.1. The condition that no four of A', B', C', D', E' be coplanar, expressed algebraically, then insures that there is a nontrivial solution (in fact, a unique nontrivial solution up to a scalar multiple) of the system, and that none of the ρ_i is zero in this solution. Substituting back, we find that (1) is determined absolutely (up to a scalar multiple). The fact that A', B', C', D' are not coplanar then shows that (2) holds.

So far, we have shown that if there is a mapping α of the form (1) that takes A, B, C, D, E into A', B', C', D', E', respectively, then we can determine it uniquely, and (2) is satisfied. But now we can show that α is, in fact, a projective transformation. The method of proof in Section 7.1 shows that the inverse image of a plane is a plane, whence the inverse image of a line (the intersection of two planes) is a line. We leave details to the reader.

So we have shown that there exists a projective transformation taking A, B, C, D, E into A', B', C', D', E'. Now, given five points A', B', C', D', E', no four coplanar, and five points A'', B'', C'', D'', E'', no four coplanar, let α be a projective transformation taking A into A', and so on, and β a projective transformation taking A into A'', and so on, We know that these exist. Then a projective transformation taking A' into A'', and so on, is $\beta\alpha^{-1}$. Thus we have proved the existence part of the theorem.

To prove uniqueness, it is enough to show that if the projective transformation α fixes five points, no four collinear, then it must be the identity. Moreover, there is no loss of generality in taking the five points to be A, B, C, D, E. We leave the verification of both these facts to the reader.

Suppose then that α leaves fixed A, B, C, D, E. Then, clearly, it maps the plane $\pi_1 = BCD$ onto itself. Let $A'(0, 1, 1, 1)$ be the point in which AE meets π_1. Then $\alpha(A')$ is a point on the image of AE and in the image plane of π_1. Since these are both invariant under α, so too is A'. Now α

induces a projective transformation of π_1 leaving fixed B, C, D, A'. We leave it to the reader to check that no three of these points are collinear. It follows, by Theorem 1, Section 3.1 of this volume, that α restricted to π_1 is the identity. In a similar manner, α restricted to $\pi_2 = CDA$ is the identity. Suppose now that X is any point of R, and that it has coordinates (x_1, x_2, x_3, x_4). Then AX meets π_1 in $X_1(0, x_2, x_3, x_4)$. Since α is the identity on π_1, X_1 is fixed under α, and since also A is fixed under α, the line $l_1 = AX$ is fixed (as a whole) by α. Similarly, $l_2 = BX$ is fixed under α. But these lines intersect in X, so that the image of X, which is the point of intersection of $l_1' = l_1$ and $l_2' = l_2$, is X itself. Thus α is the identity. ▼

9.6. Quadric Surfaces in Projective Space

A quadric surface of projective space is defined to be the set of all those points of the space whose homogeneous coordinates (in the fourth model or the first model) satisfy a homogeneous equation of the second degree:

$$a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + a_{44}x_4^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3$$
$$+ 2a_{14}x_1x_4 + 2a_{23}x_2x_3 + 2a_{24}x_2x_4 + 2a_{34}x_3x_4 = 0.$$

The analogs of the remarks of Section 8 above hold here, and by exactly the same use of the theorem of invariance of signature, we can show that there are eight projective classes of quadrics (as they are called):

(1)	$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 0,$	imaginary oval surface;
(2)	$x_1^2 + x_2^2 + x_3^2 - x_4^2 = 0,$	oval surface;
(3)	$x_1^2 + x_2^2 - x_3^2 - x_4^2 = 0,$	toroidal surface;
(4)	$x_1^2 + x_2^2 + x_3^2 = 0,$	imaginary cone;
(5)	$x_1^2 + x_2^2 - x_3^2 = 0,$	real cone;
(6)	$x_1^2 + x_2^2 = 0,$	two imaginary planes;
(7)	$x_1^2 - x_2^2 = 0,$	two real planes;
(8)	$x_1^2 = 0,$	two coincident planes.

The names of the "imaginary" surfaces are intended to be suggestive only; they become meaningful (but lose the prefix "imaginary") when we consider them as equations of surfaces in *complex* projective space.

In the following list we give a *distinguishing projective invariant* of each of the eight classes: that is, some feature invariant under projective transformations which is not shared by any of the other classes. This gives a geometric proof that the classes are projectively distinct; note, however, that it does not show that the classification is complete. Compare Volume 1, Section 31. We also give comments on some of the classes.

(1) is the empty set.

(2) has a tangent plane at every point, which meets it in that point alone ("in two imaginary lines"). The class (2) contains the projective completions of the ellipsoids, hyperboloids of two sheets, and elliptical paraboloids of affine space.

(3) has a tangent plane at every point P, which meets it in a pair of lines intersecting at P. The class (3) contains the projective completions of the hyperboloids of one sheet and the hyperbolic paraboloids.

(2) and (3) are the nondegenerate types.

(4) is a single point.

(5) has a single *singular point* (the vertex): that is, a point where there is no tangent plane. The tangent plane at every other point meets it in a line ("in two coincident lines"). The class (5) contains the cones and the conic cylinders (elliptic, parabolic, and hyperbolic). The cylinders are distinguished from the cones only in affine space: in the first model they are those cones whose center (vertex) is a special point.

(6) is a line.

(7) has a singular line (line consisting entirely of singular points). At every other point the tangent plane meets the surface in a plane.

(8) is a plane.

APPENDIX I TO CHAPTER I

The Topology of the Projective Plane

We start by saying what we mean by a topological space. Let X be a set of points. Suppose that with each point x of X we have associated a collection of subsets of X called the *neighborhoods* of x, having the following properties:

1. *x is a member of every neighborhood of x.*
2. *If u and v are two neighborhoods of x, there exists a neighborhood w of x entirely contained within the intersection of u and v.*
3. *If y is any point of a neighborhood u of x, there exists a neighborhood v of y entirely contained within u.*
4. *If x and y are distinct points, there exist disjoint neighborhoods u of x and v of y.*

We call a set X together with such a collection of neighborhoods a *topological space*; the collection of neighborhoods is called the *topology* of the topological space X.

As examples of topological spaces we have: (1) The real line—here we take the neighborhoods of each point x to be the open intervals (intervals without their end points) containing x. (2) The unit circle—here we take the neighborhoods of x to be the open arcs containing x. (3) The plane—we take the neighborhoods of x to be the open disks (interiors of circles) with center x.

We leave it to the reader to verify that all these are topological spaces according to our definition.

We can now define a continuous mapping of one topological space onto another. Let X and Y be topological spaces, and let α be a mapping of X onto Y. Given a point y of Y, suppose that for every neighborhood v of y and every x in X such that $\alpha(x) = y$ we can find a neighborhood u of x such that α maps u entirely into v. Then α is said to be *continuous at* y. If α is continuous at every point of Y, it is said to be *continuous*. Intuitively, the reader can think of a continuous mapping as one which may stretch X but will not tear it—points sufficiently close together in X will be taken into points not too far apart in Y.

Suppose now that α is a one-one mapping of X onto Y such that both α and the inverse mapping α^{-1} are continuous. Then we say that α is a *homeomorphism*, and that X and Y are *homeomorphic*. Topology is the study of those properties of spaces which are preserved under homeomorphisms; such properties are called *topological* properties. In this chapter we study the topological properties of the projective plane. Of course, we first have to say what our neighborhoods are to be. We start, however, by considering the topology of the projective line.

The projective line L (first model) is constructed from the ordinary line l by the addition of one ideal point. The reader can check that we can construct models (2), (3), and (4) for it; (2) will be the rays in the plane through the origin S, (3) will be identified pairs of antipodal points on the unit circle, and (4) will be the collection of all pairs (x_1, x_2) of real numbers, with proportionate pairs identified. Of course, the projective line is not as interesting an object as the projective plane, but some things can be done (for example, we can define cross ratio and define projective mappings as those which preserve it).

To make a topological space of L, we define (in the second model) a neighborhood of the ray x to be the set of all those rays lying within some open angle containing the ray x ("open"

means without the bounding rays). In the first model, the definition does not turn out to be so elegant, but it is the natural one when we consider the correspondence between the first and second models. A neighborhood of an ordinary point of L is a neighborhood when considered as a point of l (that is, an open interval containing x); a neighborhood of the ideal point is the *complement* in L of any closed segment of l (Fig. 41). The reader should check that the definitions in both models do give us a topological space and that they "tie in" together. He should also consider how we should define the neighborhoods of a point in the third model.

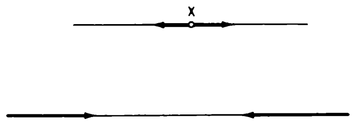

Fig. 41

It turns out now that with this topology the projective line is homeomorphic to a circle (where the circle has the topology we described earlier). To see this, let us construct a circle C touching L (in the first model) at the point S, and let N be the point of C antipodal to S. We make correspond to each point x of C the point y of L in which Nx meets L, and we put N itself in correspondence with the special point. It is easy to see (Figs. 42 and 43) that this is a homeomorphism.

The topological equivalence that we have established between the projective line and the circle allows us to say that a projective line is a closed curve of the projective plane (that is, that it has no end points).

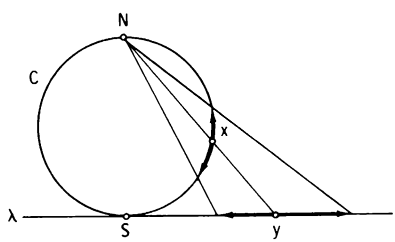

Fig. 42

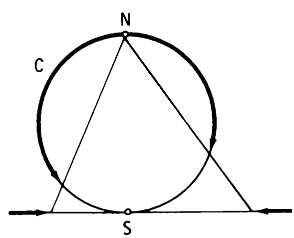

Fig. 43

A Euclidean line l is divided into two by every point x of it; more precisely, the line l minus a point x is the union of two subsets such that every point of one of the subsets has some (sufficiently small) neighborhood not meeting the other. Such subsets are called *disconnected*, and we may say that the removal of x *disconnects* the line.

On the other hand, the projective line does not have this property: there is no point of the line the removal of which would disconnect it. For this is exactly the situation on the circle, and notions of connection are topological.

Furthermore, the removal of distinct points x_1, x_2 from the Euclidean line disconnects it into three components: $(-\infty, x_1)$, (x_1, x_2), (x_2, ∞) (where we may suppose $x_1 < x_2$). On the other hand, the removal of two points divides the circle (and hence also the projective line) into *two* components A and B. Each component is homeomorphic to an interval, and the adjunction of x_1 and x_2 to either A or B produces a set homeomorphic to a line segment (closed interval). In other words, two points x_1 and x_2 determine *two* line segments, not one, and, in specifying a segment of a projective line, we have to give a third point (lying on the segment) to make it clear which of the two we have in mind.

Let us now pass on to an examination of the topological properties of the projective plane.

We define a neighborhood of a point x to be the set of all the interior points of some oval curve C such that x is an interior point of C (see Section 9.4 of this volume).

We note some properties of the projective plane that distinguish it from the Euclidean plane. If A is a topological space, we say that a subset B of A is a *component* of A provided B is connected but no subset of A properly containing B is connected. Now if A is what is left of the Euclidean plane π after the removal of a line l, then A consists of two components. By analogy with what we said for the line above, we may thus say that l (or more strictly the removal of l) disconnects the Euclidean plane into two components.

In the projective plane, this is false; no line disconnects it. For let L be a line of Π, and let x and y be any two points of Π

not on L. We wish to show that x and y belong to the same component, for if we can prove that this holds for *every* pair of points, we shall know that the whole of Π minus L can have only one component. Now the line xy meets l in a point P. Consider the segment $[x, y]$ that does not contain P. It is certainly connected (being homeomorphic to a segment of the real line), so that x and y must belong to the same component.

Two intersecting lines of Euclidean space disconnect the segment into four components, each a domain (a domain is a connected set that contains a neighborhood of every point in it). On the other hand, two projective lines disconnect the projective plane into two domains. To see this, we consider cases, using the first model.

Case 1. Let L and M be the lines, and suppose that they intersect in an ordinary point X. We wish to show that Π minus the points of L and M consists of two disjoint domains. We take A to consist of all the (ordinary) points lying in one of the pairs of vertically opposite angles formed by L and M, together with the ideal points of the lines in them and through X. We define B similarly. Let us show that A is a domain. If x is an ordinary point of A, then a sufficiently small circle with center x will lie in A and will be an oval curve of Π (since the "ellipses" in the first model consist entirely of ordinary points). Thus it will be a neighborhood of x lying in A. If x is an ideal point of A, we take the neighborhood of x to be the interior of any hyperbola with asymptotes L and M (and together with its ideal points), such that x lies in its interior. It is clear that such "hyperbolas" exist.

We have shown that A is open; that is, it contains a neighborhood of each of its points. To prove that it is connected, let x and y be any two points of A. Then one of the two segments determined by x and y lies entirely in A, as the reader may verify by sketching a few cases. We have thus shown that A is a domain, and we may show similarly that B is a domain. It remains only to show that A and B, together with the lines L and M, exhaust all of Π. It is clear that they account for all the ordinary points of Π, and if x is an ideal point not on L or

M, the line xX lies entirely within A or B (every line through X does), and x belongs to A or B accordingly.

Case 2. The lines L and M are both ordinary, but they meet on the ideal line. In this case we take A to consist of all the ordinary points lying between L and M, and B to consist of everything else (except the points of L and M themselves, of course). We leave the proof that A and B are both domains to the reader.

Case 3. One of L and M, say M, is ideal. In this case we take A to be all the points on one side of L, and B all the points on the other. It is clear that A and B are open and connected, and are therefore domains, and it may be checked that they are components.

Three nonconcurrent lines of the Euclidean plane divide it into seven domains (Fig. 44). On the other hand, three noncurrent lines L, M, N of the projective plane divide it into four domains (Fig. 45).

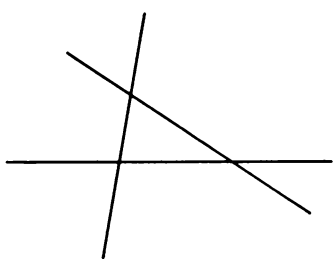

Fig. 44

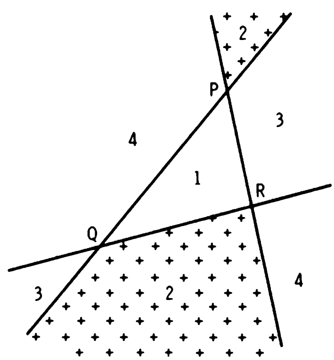

Fig. 45

We consider only the case where L, M, and N are ordinary lines meeting in the ordinary points P, Q, R (Fig. 45). Then one of the four domains is the interior of triangle PQR. The second of the four domains, numbered 2, is marked off by

crosses in the figure, and consists of three parts: (1) all the interior points of the angle at P opposite to the angle QPR; (2) all the interior points of the angle QPR, except for those of the triangle PQR and its side QR; (3) all the ideal points of the lines through P and passing within the angle QPR.

We define domains (3) and (4) similarly. We now show that they really are domains.

I. The set (1) is a domain. For each point of it may be surrounded by a circle lying entirely inside (1), and this is a neighborhood. Thus (1) is open. If x and y are two points of it, the (ordinary) segment joining them is connected, so that (1) must be connected.

II. Consider the set (2) (the proofs for (3) and (4) will be parallel). If x is an ordinary point, it has a (sufficiently small) circular neighborhood. If x is ideal, we may find a hyperbola having asymptotes PQ and PR and not intersecting QR but lying within 2. This hyperbola, together with the ideal points of its asymptotes, is an oval curve of Π lying within (2) and having x as an interior point. Thus (2) is open. If x and y are distinct points of (2), one of the segments they determine lies within it, as may easily be verified. So (2) is connected and thus is a domain.

We recommend the reader to complete the proof by considering the other possible cases—when L is the ideal line and when L and M meet in an ideal point.

Thus, if we take three noncollinear points of the projective plane and draw the lines through pairs of them, we obtain not one triangle but four. In order to distinguish them, we need to specify one more point lying inside them or say which of the two segments PQ, which of the two segments QR, and which of the two segments RP are meant to be in its sides. Note that we are talking of triangles as not simply sets of three points and three lines but as domains of the plane bounded by such. The Euclidean triangle may be regarded as a projective triangle, embedded in the ordinary part of the projective plane, that is, as a triangle having no ideal points in its interior or on

its boundaries. Note that there exist projective triangles having no ideal points in their interior but having ideal boundary points (of course, we are working the whole time within the first model). The reader should think of an example of one.

We come now to the fundamental result of this appendix:

Theorem. *A projective plane is a nonorientable two-dimensional closed manifold of Euler characteristic* 1. *Conversely, any space having all these properties is homeomorphic to the projective plane.*

Thus, the properties listed characterize the projective plane topologically (that is, up to homeomorphism).

We shall not give a proof of this theorem and shall devote the rest of this chapter merely to explaining exactly what all these properties mean.

A topological space R is called a *2-manifold* (that is, a two-dimensional manifold) if it is connected and if every point has a neighborhood homeomorphic to the interior of a circle (an open disk). A topological space is called *sequentially compact*, provided that every infinite subset has a limit point (that is, a point in R every neighborhood of which meets the subset). A 2-manifold which is sequentially compact is called *closed* (sometimes the word "closed" is used only if, in addition, the space has no boundary, but here this property follows from the others).

We show first that the projective plane is a 2-manifold, that is, that every point has a neighborhood homeomorphic to the open disk. If x is an ordinary point, we may take as a neighborhood the interior of any circle center x. If x is ideal, there exists a hyperbola (a projective hyperbola: that is, an ordinary one with the ideal points of its asymptotes added) such that x lies in its interior. We take the neighborhood of x to be the interior points of the hyperbola, and we must show now that this is homeomorphic to the open disk.

Now we may find a projective transformation of the plane that takes the hyperbola into a circle, and it will take interior

points into interior points (Section 8.4 of this volume). The inverse of this transformation is also a projective transformation. So in order to prove the homeomorphism, we need merely show that projective transformations are homeomorphisms of the projective plane.

Certainly a projective transformation is one-one and onto. We need only show its continuity, since the inverse transformation, being also projective, will then also be continuous. Suppose then that x is any point of Π, and $y = \alpha(x)$ is its image under the projective transformation α. Let $v(y)$ be any neighborhood of y, that is, the set of all interior points of some oval curve of the second degree containing y in its interior. Then α^{-1}, being a projective transformation, maps the oval curve into an oval curve, and the interior into the interior, so that it maps the neighborhood v of y onto a neighborhood u of x. Clearly, α maps u into v. This shows that α is continuous.

We have thus shown that every point of Π has a neighborhood homeomorphic to an open disk, so that Π is a 2-manifold.

We show next that Π is a connected topological space. It is certainly connected, for any two points x, y lie in a connected subset, for instance, either of the line segments joining them. We may check all the properties (1) through (5) very easily when we take x and y to be ordinary points; if one, or both, of them is ideal, let α be a projective transformation taking them both into ordinary points. We may then perform the necessary construction on $\alpha(x)$ and $\alpha(y)$, and α^{-1} will map the neighborhood or neighborhoods we obtain homeomorphically onto neighborhoods of x and y, and these neighborhoods will have the required properties.

We show next that the projective plane is sequentially compact. To do this we use the third model; the points of Π we take to be pairs of antipodal points on the unit sphere. A neighborhood of a "point" is any pair of antipodal "caps" —the interiors of circles described on the surface of the sphere. With this definition of neighborhood in the third model, it is

easy to check that the first and third models are homeomorphic, and since sequential compactness is a topological property, it is enough to prove it in the third model. Now the unit sphere in space is certainly sequentially compact, being a closed and bounded subset of Euclidean space. If A is any infinite subset of Π, we may consider it as a subset of the surface of the sphere in ordinary space. Suppose that M is a limit point of A. Then the projective point consisting of M and the antipodal point is a limit point for A thought of as a subset of Π.

We have now shown that the projective plane is a closed 2-manifold, or a *surface.*

In order to explain what we mean by saying that the projective plane is nonorientable, we must first introduce the important concept of a *triangulation.*

We will call a collection of triangles in Euclidean n-space a *triangulation*, provided that two triangles of the set have either exactly one vertex in common, exactly one side in common, or nothing in common. The set of points belonging to the triangles of a triangulation is called a *polyhedron.* A homeomorphic image of a polyhedron is called a *curvilinear polyhedron*, and the partition of the curvilinear polyhedron induced by the given triangulation of the polyhedron is called a *triangulation* of the curvilinear polyhedron.

Earlier we split the projective plane into four triangles (Fig. 45). However, this division is not a triangulation of Π, since each pair of triangles has three vertices (and one side) in common. To obtain a triangulation of the projective plane, choose one of the triangles ABC, and in it inscribe a triangle DEF, so that D lies on the side (segment) BC of ABC, E on CA, and F on AB. Next, draw those segments AD, BE, CF that do not pass through our triangle DEF.

We then obtain a triangulation of Π into the following triangles: DEF; AEF, BFD, CDE; DAB, DAC, EBC, EBA, FCA, FCB. We have to specify which of four possible triangles each of these denotes. DEF is already fixed. In all the others we have already fixed one side (EF, FD, DE, in the first three; DA, EB, FC in the last six). Each of the other sides has vertices

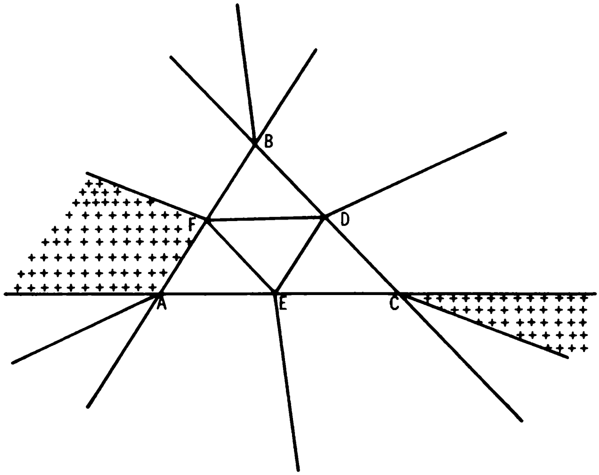

Fig. 46

on a line on which we have marked a third point, and we choose that segment which does not contain it. For example, in the triangle FCA (Fig. 46) the segment FC is already fixed, and we choose the segment FA that does not contain B, and the segment AC that does not contain E.

In Fig. 46 we show the appearance of Π in the first model when we choose DEF to lie inside the *bounded* triangle ABC. The reader should sketch for himself the figure we obtain by making DEF lie in one of the other triangles ABC. It is also of interest to sketch the triangulation we obtain in this way when BC (say) is the ideal line. In every case, the reader will find that the rules that we gave for determining the triangulation still are meaningful and completely determine the triangulation.

The reader may check directly that this collection—six points, fifteen segments and ten triangles—is a triangulation. It is proved in topology that in Euclidean 4-space we can construct a polyhedron with six vertices, fifteen edges, and ten triangular faces such that this polyhedron is homeomorphic

to the projective plane by a homeomorphism that induces the given triangulation. This shows directly that what we have is indeed a triangulation.

There exist many other triangulations of the projective plane. There is even one simpler than the given one, with five points, twelve segments, and eight triangles. It is an interesting exercise to construct this triangulation given no further information.

We now consider *orientation*. A triangle, together with a specified direction in which it is described, is called *oriented* (see Volume 1, Section 6). The two possible directions in which the triangle may be described are called *orientations* of the triangle. Thus we may say the triangle ABC has the two orientations $\overrightarrow{ABC} = \overrightarrow{BCA} = \overrightarrow{CAB}$ and $\overrightarrow{ACB} = \overrightarrow{CBA} = \overrightarrow{BAC}$. Each of the orientations determines a definite direction along each side. Thus the first of our orientations for ABC determines the directions $\overrightarrow{BC}$, $\overrightarrow{CA}$, $\overrightarrow{AB}$ on the sides BC, CA, AB, respectively, of the triangle. Such a direction along a side is called an *orientation* of the side.

Suppose now that we have given an orientation to every triangle of some triangulation K. Let P and Q be two triangles of K that have a side in common. If the given orientations of P and Q induce opposite orientations on the common side, we say that P and Q have the *same orientation*. Conversely, if P and Q induce the same orientation on their common side, we shall say they have *opposite orientation*. For example, in Fig. 47, we show two oriented triangles $\overrightarrow{ABC}$ and $\overrightarrow{ABD}$, having the common side AB. Since they induce the same orientation $\overrightarrow{AB}$ on AB, they have opposite orientation.

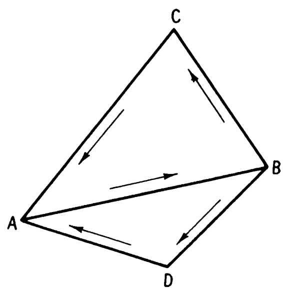

Fig. 47

Definition. A surface is said to be *orientable*, provided that there exists a triangulation of it such that the triangles of

the triangulation can be assigned orientations in such a way that any two adjacent triangles have the same orientation. If such a triangulation does not exist, we say the surface is *nonorientable*.

Since a homeomorphic image of a triangulation is again a triangulation, and that of an oriented triangle is again an oriented triangle, it is clear that a homeomorphic image of an orientable space is orientable and that a homeomorphic image of a nonorientable space is nonorientable. Thus orientability is a topological property. It may be proved that if a space has one "orientable triangulation" (in the sense of the definition above), then every triangulation is orientable, and, conversely, if there is one nonorientable triangulation, then every triangulation is nonorientable. We may even refine this result, as follows.

Let $\Delta_1, \Delta_2, \ldots, \Delta_n$ be a finite sequence of oriented triangles of some triangulation, such that each successive pair Δ_i, Δ_{i+1} and also Δ_n, Δ_1 have a side in common. We say the sequence $\{\Delta_i\}$ is *disorienting* provided that each pair of neighboring triangles has the same orientation, yet Δ_n and Δ_1 have opposite orientations. Then a surface is nonorientable if and only if there exists a disorienting sequence on it.

As examples of orientable surfaces, we have the sphere, torus, pretzel, square, surface of a cylinder, and many others.

An example of a nonorientable surface is provided by the Möbius band. It may be obtained by sticking together opposite sides of a rectangular ribbon, in such a way that points symmetric with respect to the center of the rectangle are made to coincide. To do this we simply give one twist to the ribbon before sticking it together (Fig. 48). Let us show that the Möbius band is nonorientable.

Consider the triangulation of the Möbius band given in Fig. 49. Here C is the same point as B, and D as A. Thus the first and the last triangle induce the same orientation $\overrightarrow{BA}$ on their common side AB. The reader can check in a glance that the triangles of the triangulation do, in fact, satisfy the conditions for a triangulation.

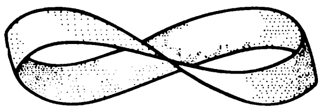

Fig. 48

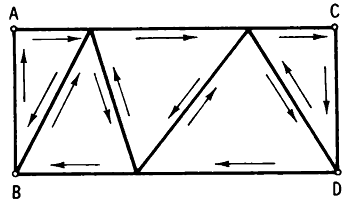

Fig. 49

In order to show that the projective plane is nonorientable, we shall construct a fifth (and final!) model, containing a copy of the Möbius band.

We start with the third model and throw away the whole southern hemisphere but keep the equator. Then a point of the fifth model will be an ordinary point of the sphere if it is not on the equator and a pair of antipodal points if one of them does lie on the equator. In effect, we have made the third model more economical, by making most of the "points" consist of one ordinary point instead of two, and we have done it in the very simple manner of simply throwing out one member of each (nonequatorial) pair. We take a neighborhood of a nonequatorial point to be the open cap bounded by any circle on the surface of the sphere and a neighborhood of an equatorial "point" to be a pair of opposite half caps. It is clear that the third and fifth models are homeomorphic and also projectively equivalent under the mapping that sends each "point" of the third model into its remaining representative in the fifth (or to both of the equatorial "points"). The lines in the fifth model will be halves of great circles, with their end points identified.

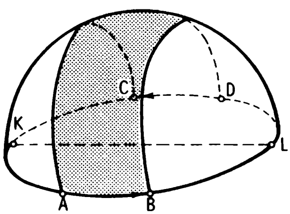

Fig. 50

In this model, let us cut out the strip lying between two parallel planes, each perpendicular to the equatorial plane (Fig. 50). This intersects the projective plane in a rectangular strip, except that points on the arc in which the strip meets the equator have to be identified

with the antipodal points. This means that we have to identify the edge AB and the edge DC (Fig. 50), so that our strip is just a copy, in the projective plane, of the Möbius band.

We can now immediately show that the projective plane is nonorientable. For triangulate this Möbius band as in Fig. 49, and extend this triangulation to a triangulation of the whole space. Then the Möbius band is a disorienting sequence of this triangulation, and the result follows. Since orientability is a topological property, all the other models (as well as any new ones we might think of) for the projective plane are also nonorientable.

It is interesting to examine what is left of the projective plane Π after we have removed the Möbius strip. To see what we get, we have to stick the remaining two surfaces along their bottom rims, so that D coincides with A, L with K, and B with C. This may actually be done if we first take a mirror image of one of the pieces in the equatorial plane. If we then twist the reflection round and translate it until it is under the other piece, and stick it on, we obtain a hemisphere (these operations preserve all topological properties). Since a hemisphere is homeomorphic to a disk, we may say that Π is obtained by sticking a disk and a Möbius band together along their edges. This sticking together cannot be accomplished in Euclidean 3-space (ordinary space) unless we are prepared to allow one surface to pass through another, imagining that there is no tearing. However, it may be accomplished in 4-space, as follows.

Place a Möbius band M in a three-dimensional subspace of 4-space (we may imagine it as the xyz "hyperplane"); let S be any point not in this subspace [e.g., $(0, 0, 0, 1)$, on the fourth axis]. Draw all the segments from S to points on the boundary of M. Since this boundary is homeomorphic to the circumference of a circle, the surface developed by these segments is homeomorphic to a disk. Since, moreover, each line through S meets the subspace in which M lies only in the point where it meets the boundary of M, the surface we have constructed does not intersect itself. To complete the triangulation of the Möbius strip to a triangulation of all of Π (as we said we could earlier),

it is enough to join S to those edges of the triangles in the triangulation of M that lie on the boundary of M. This provides a triangulation in Euclidean 4-space which is evidently homeomorphic to this model of Π in 4-space, and therefore to any projective plane. The construction we have given is due to V. A. Yefremovich.

It remains to explain what we mean by the Euler characteristic of a surface.

Let T be a triangulation of a surface S, and let V be the number of vertices of T, E the number of edges, and F the number of faces. Then we define the *Euler characteristic* $\chi = \chi(S)$ of the surface by

$$\chi = V - E + F.$$

This concept is useful only because it is independent of the triangulation T that we use. Furthermore, the Euler characteristic is invariant under homeomorphism, and so is a topological invariant. We consider a number of examples.

Example 1. *The sphere.* Let us inscribe in the sphere S a regular octahedron. The octahedron is certainly a triangulation in 3-space, and we obtain a triangulation of S by projecting H radially into S from its center. In this triangulation (Fig. 51), we have

$$\chi(S) = V - E + F = 6 - 12 + 8 = 2.$$

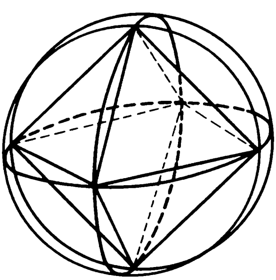

Fig. 51

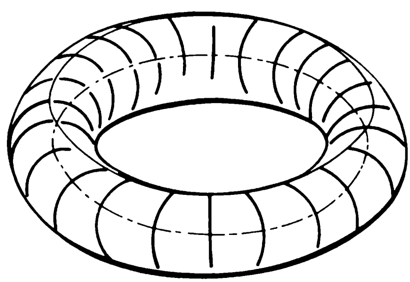

Fig. 52

Example 2. *The torus T* (Fig. 52). We construct a polyhedron homeomorphic to T as follows: We arrange nine identical cubical

blocks in a square and remove the middle one (Fig. 53). We triangulate the remainder as shown in Fig. 54.

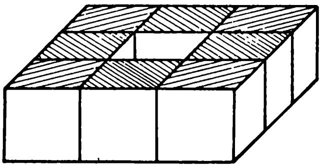

Fig. 53

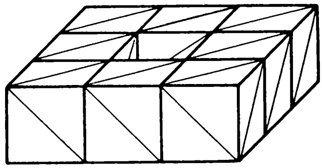

Fig. 54

Since our figure is homeomorphic to the torus, we have found a triangulation for T and have

$$\chi(T) = V - E + F = 32 - 96 + 64 = 0.$$

Example 3. *The vertical face C of a circular cylinder.* We inscribe in C a triangular prism, each face of which we divide by a diagonal into two triangles. The triangulation we obtain is homeomorphic to the cylinder under horizontal projection through the axis of the cylinder. Thus we have found a triangulation for C and have

$$\chi(C) = V - E + F = 6 - 12 + 6 = 0.$$

The reader should check for himself that we get the same value for χ if we start with a prism whose base is a regular n-gon.

Example 4. *The disk D.* This is homeomorphic to a triangle, for which

$$\chi(D) = V - E + F = 3 - 3 + 1 = 1.$$

Alternatively, it is homeomorphic to a regular n-gon, which we may triangulate by drawing all the lines from the vertices to the center. For this triangulation we find that

$$\chi(D) = (n + 1) - 2n + n = 1.$$

Example 5. *The Möbius band M.* We triangulate M as in Fig. 49 and find that

$$\chi(M) = 5 - 10 + 5 = 0.$$

The reader should see what happens when we take the triangulation

in which there are n division points along BD and either n or $n+1$ along AC. Why can we not take 0 and 1 or 0 and 1 here?

Example 6. *The projective plane* Π. To calculate its Euler characteristic, we use the triangulation given in Fig. 46. We have

$$\chi(\Pi) = 6 - 15 + 10 = 1.$$

We have given a number of properties of the projective plane, which characterize it up to homeomorphism. We can regard our properties as providing a topologically complete description of Π; any further topological properties must be deducible from these.

As we said, we will not give a proof of the theorem in this book.

CHAPTER II

Inversion

10. The Power of a Point with Respect to a Circle

The *power* of a point M with respect to a circle C in the Euclidean plane is defined to be the number

$$\sigma(M) = \overrightarrow{MA} \cdot \overrightarrow{MB},$$

where A and B are the points of intersection of MO and C, and O is the center of C. By a famous classical theorem, we have

$$\sigma(M) = \overrightarrow{MA'} \cdot \overrightarrow{MB'},$$

where A' and B' are the points of intersection with C of *any* line l through M. In particular, if M lies outside C, σ is the square of the length of a tangent from M to C.

It is clear that σ is positive when M lies outside C (since, in this case, the vectors $\overrightarrow{MA}$ and $\overrightarrow{MB}$ have the same sense), negative if M lies within C (since now the vectors have opposite senses), and zero if M lies on C. We may regard σ as a measure of how far M lies from the circumference of C, and it is the most natural such measure for a number of reasons. We give only one such reason here.

Suppose that we introduce rectangular Cartesian coordinates. Then any circle C will have an equation of the form

$$ax^2 + ay^2 + 2gx + 2fy + c = 0.$$

It is not hard to prove that the power of the point $M(x_0, y_0)$ with respect to C is then precisely

$$\sigma(M) = ax_0{}^2 + ay_0{}^2 + 2gx_0 + 2fy_0 + c.$$

We may also show that if C has radius r and $MO = d$, then

$$\sigma = d^2 - r^2.$$

For if M lies outside C, we already know that σ is the square of the distance from M to C along a tangent line. By an easy application of the Pythagorean theorem, we see that this equals precisely $d^2 - r^2$.

If M lies within C and MO meets C in A while OM meets C in B, then

$$\begin{aligned}\sigma &= \overrightarrow{MA} \cdot \overrightarrow{MB} = (\overrightarrow{MO} + \overrightarrow{OA})(\overrightarrow{MO} + \overrightarrow{OB}) \\ &= (\overrightarrow{MO} - \overrightarrow{AO})(\overrightarrow{MO} + \overrightarrow{OB}) \\ &= (\overrightarrow{MO} - \overrightarrow{AO})(\overrightarrow{MO} + \overrightarrow{AO}) = MO^2 - AO^2 = d^2 - r^2.\end{aligned}$$

Incidentally, this proof also holds when M lies outside C, and may be used as a proof of the Pythagorean theorem.

Finally, if M lies on C, the result is clear, since $d = r$.

II. Definition of Inversion

Let C, as before, be the circle with center O and radius r. We define a transformation of the plane (excluding the point O) by making correspond to each point M the point M' on the line OM such that $\overrightarrow{OM} \cdot \overrightarrow{OM'} = r^2$.

It may be checked that this is a one-one transformation of the whole plane (excluding the point O) onto itself. We call it an *inversion* with *pole* O and *coefficient* r, and we write it (O, r). We note that if (O, r) sends M into M', then it sends M' back into M, so that *an inversion is an involution* (its square is the identity).

If I is the inversion (O, r), then the points of the circle C with center O and radius r remain fixed under I, and they are the

only fixed points of I. In general, we may give a simple geometric construction for the inverse of a point with respect to C (this is an alternate way of saying "the image of a point under the inversion I with respect to the circle C").

1. If M lies outside C, let one of the tangents from M to C meet it in T. Then M' is the projection of T onto OM.

2. If M lies inside T, we reverse the construction. Let T be one of the points in which the perpendicular to OM through M meets C. Then the tangent to C through T meets OM in the required point M'.

If M lies on C, both these constructions come to the same thing and give us $M' = M$.

To prove that these constructions give the required point, suppose, for instance, that M lies outside C and that M' lies inside C. Let a tangent from M meet C in T. Then the triangles OTM' and OMT are similar (since they have the same angle at O and are both right-angled). Thus

$$OM : OT = OT : OM',$$

or

$$OM \cdot OM' = OT^2 = r^2,$$

so that M and M' are images of each other under I.

We see that the inverse with respect to C of a point inside C is a point outside C, and that, conversely, the image of a point outside C is a point inside C. As we said before, the points of C are their own images.

12. Properties of Inversion

Theorem 1. *The images under $I = (O, r)$ of four points on a cir 'e S not through O are four points of a circle S' not through O.*

Proof. Let the four given points of S be A, B, C, D, and let A', B', C', D' be their images under I. Then

$$\overrightarrow{OA} \cdot \overrightarrow{OA'} = \overrightarrow{OB} \cdot \overrightarrow{OB'} = \overrightarrow{OC} \cdot \overrightarrow{OC'} = \overrightarrow{OD} \cdot \overrightarrow{OD'} = r^2. \quad (1)$$

Let A_0 be the second point in which OA meets S (if OA is

tangent to S, we set $A_0 = A$). Define B_0, C_0, D_0 similarly. Let σ be the power of O with respect to the circle S. Then

$$\overrightarrow{OA} \cdot \overrightarrow{OA_0} = \overrightarrow{OB} \cdot \overrightarrow{OB_0} = \overrightarrow{OC} \cdot \overrightarrow{OC_0} = \overrightarrow{OD} \cdot \overrightarrow{OD_0} = \sigma. \quad (2)$$

It follows from (1) and (2) that

$$\overrightarrow{OA'} : \overrightarrow{OA_0} = \overrightarrow{OB'} : \overrightarrow{OB_0} = \overrightarrow{OC'} : \overrightarrow{OC_0} = \overrightarrow{OD'} : \overrightarrow{OD_0} = r^2 : \sigma. \quad (3)$$

Note that division by σ is legitimate, for if $\sigma = 0$, O would lie on S. Now (3) states that $A' = \gamma(A_0)$, where γ is the homothetic transformation with center O and coefficient r^2/σ (see Volume 1, Section 16)—and similarly for B', C', D'. Since γ takes circles into circles, A', B', C', D' will lie on the image circle S' of S under γ. ▼

Corollary 1. *Inversion with respect to a circle C induces a one-one mapping of any circle S, not through the center O of C, onto a circle S', also not through O.*

Proof. Let S' be the circle through the images A', B', C', D' of four given points A, B, C, D of S. We first show that S' does not pass through O. Let D be a large circle with center O that contains all of S inside it. Then the image of D under I will clearly be a small circle D' with center O. Now I takes every point inside D into a point outside D', for the inverse of a number smaller than the radius of D is a number larger than the radius of D'. It follows that S' lies entirely outside D' and so cannot pass through O.

We have shown that every point of S is mapped into a point of S'. Conversely, the inverse I of I maps S' into a circle (by what we have already proved), which must clearly be S. So I is both one-one and onto. ▼

Corollary 2. *If S has radius R and its image S' under $I = (O,r)$ has radius R', then*

$$R'/R = r^2/\sigma,$$

where σ is the power of O with respect to S.

Theorem 2. *Under an inversion I, the images of four points of a circle S through the center O of I are four points of a line l not through O.*

Corollary 1. *Inversion maps a circle through the center of inversion (all except for the center itself, of course) one-one onto a line not through the center.*

Proof. It is clear that it is enough to prove the corollary. Suppose then that S is a circle through the center O of I. Let the point of S antipodal to O be E, and let E' (on OE) be the image of E under the inversion. Then the perpendicular l to OE through E' is the image of S under I.

For let A be any point of S. If $A = E$, then $A' = E'$ lies on l. If not, let OA meet l in A'. Since $A \neq O$ these lines cannot be parallel. We assert that A' is the image of A under I. For the angles OAE and $OE'A'$ are right angles (the first is the angle in a semicircle, and the second is a right angle by construction). It is then easy to see that the triangles $OE'A'$ and OAE are similar, so that

$$OA \cdot OA' = OE \cdot OE' = r^2.$$

We have thus shown that I maps S into l. But exactly the same argument shows that $I^{-1} = I$ maps l into S. So I maps S onto l, and it is clearly one-one. ▼

Corollary 2. *Under inversion, the image of a line l that is not through the pole O of I is mapped one-one onto a circle S through O.*

Theorem 3. *The image under inversion of a line through the pole of the inversion (all except for the pole itself) is the line (less the pole) itself.*

Let us say that two circles S and S' *meet at an angle* α, provided that the tangents to them at their points of intersection

make angles α (the pairs of tangents at the two points of intersection will clearly make the same angles). Let us also say that the circle S and the line l *meet at the angle* α, if the tangents to S at the points in which l meets it make angles α with l. We already know what is meant by the angle between two lines. We may regard a tangent to a circle as a line meeting it at an angle zero and, similarly, two circles tangent to each other as circles meeting each other at an angle zero. It will be convenient to regard parallel lines as meeting each other at an angle zero. Although the lines do not, in fact, meet, this is a natural convention. For reasons that will appear in the next section, let us agree to regard all lines and all circles as being *conformal circles*. Then we have:

Theorem 4. *The angles between intersecting conformal circles are preserved under inversion.*

This means that if S_1 and S_2 are two intersecting conformal circles and S_1', S_2' are their images under an inversion I, then the angle between S_1 and S_2 is the same as the angle between S_1' and S_2'.

Proof. Suppose first that S_1 is an ordinary circle and that S_2 is a line l tangent to it. We distinguish three cases:

Case 1. S_1 does not pass through O. In this case, the images of S_1 and S_2 will be two circles (in case l does not pass through O) or a circle and a line (if l does). In either case, since S_1 and S_2 have only one point in common, S_1' and S_2' will also have only one point in common, which can only happen if they are tangent to each other.

Case 2. S_1 passes through O, but S_2 does not. In this case, the images are a circle and a line, and the same argument holds.

Case 3. S_2 is tangent to S_1 at O. Then the images will be parallel lines, which touch by convention.

Suppose now that S_1 and S_2 are any conformal circles, meeting in a point M. If M is the only point in which they meet, and $M = O$, then the simple proof is left to the reader. In all other cases we may choose a point M of intersection of S_1 and S_2 that is not O. Let l_1 be the tangent line to S_1 at M, and let l_2 be the tangent line to S_2 at M (if one or both of S_1 and S_2 are lines, we understand the tangent line to be the line itself). Let S_1', S_2', l_1', l_2' be the images of these conformal circles under I. By what we have already proved, l_1' is tangent to S_1' and l_2' is tangent to S_2'. This means (if M' is the image of M) that S_1' and S_2' will meet in M' at the same angle as l_1' and l_2' meet in M'. So it is enough to prove the following special case:

If l_1 and l_2 are lines meeting at an angle α in the point M, and $O \neq M$, then the images l_1' and l_2' of l_1 and l_2, under an inversion with center O, are conformal circles also meeting at an angle α. We distinguish cases:

1. l_1 passes through O. Let m be the line through O perpendicular to l_2. Then $l_1' = l_1$, and l_2' is a circle through O with a diameter on m. It follows that the tangent to l_2' at O is the perpendicular to m through O and so is parallel to l_2. So the angle between l_2' and $l_1' = l_1$ is the same as the angle between l_2 and l_1.
2. Neither of l_1 and l_2 passes through O. Let m_1 and m_2 be the perpendiculars from O onto l_1 and l_2, respectively. Then l_1' is a circle through O with a diameter on m_1, and l_2' is a circle through O with a diameter on m_2. It follows that the tangents to l_1' and l_2' at O will be perpendicular to m_1 and m_2, respectively, and so parallel to l_1 and l_2, respectively. So the angle in which they meet at O is equal to the angle between l_1 and l_2. ▼

Note. It can be shown that under inversion the angles between *any* two intersecting curves are preserved. Here by "curve" we mean curve with a tangent at every point; the angle between two intersecting curves (at a point M of intersection) is defined as the angle between their respective tangents at M, and we say this angle is preserved under I if the angle

between the image curves at the point M' is the same as the angle between the given curves at M. Transformations having this "angle-preserving" property are called *conformal.* Thus inversion is a conformal transformation of the plane (less one point).

13. Circle Transformations and the Fundamental Theorem

So far, we have had to exclude the center of an inversion from the plane whenever we wished to consider an inversion as a transformation. This is a nuisance, and the whole theory becomes tidier when we remove this singularity. We adjoin to the plane one more point, called the *special point* ω, and which the reader may think of as being infinitely far away, but in no particular direction. The new set we obtain (all the points of the plane, together with the special point) is called the *conformal plane*, or the *Riemann sphere.*

The reason for the latter name is as follows: Let Σ be a sphere lying on the Euclidean plane E, and touching it at the south pole S. Let N be the north pole (the point antipodal to S), and let α be the mapping that sends each point M of Σ onto the point M' in which the line NM meets E. This is a one-one mapping onto all of E, but it is not defined on all of Σ; it is defined everywhere except at the point N. By adjoining a special point to E, we may extend α to a one-one mapping of all of Σ onto all of the conformal plane. Thus the Riemann sphere may be thought of as a punctured sphere stretched flat, with the puncture point then restored.

Let I be any inversion of the plane E. Then we may extend the definition of I to all of the conformal plane K (obtained from E by adjoining the special point) by setting $I(O) = \omega$, $I(\omega) = O$. On K, I is now a one-one mapping of all of K onto all of K.

If F is the name of any figure of E, let us agree to give the name "conformal F" to a figure T of K as follows:

If F is a bounded figure (that is, contained inside some

sufficiently large circle), then T can be called "conformal F" if it consists of the same points as F.

If F is unbounded, then T can be called "conformal F" if it consists of the points of F together with the special point.

These definitions may appear a little strange, but they are quite natural. We did exactly the same in the projective plane when we talked of projective hyperbolas, for example, meaning ordinary hyperbolas with the addition of the appropriate ideal points. What do we obtain here? A conformal circle will be just an ordinary circle, since a circle is a bounded figure. On the other hand, a conformal line will be an ordinary line together with the special point. Let us refer to lines (in this sense) and circles together as "conformal circles." Thus the straight lines of K are just those conformal circles that pass through the special point. Any two intersecting conformal circles either touch or intersect each other in two points. If they are both ordinary circles, this is clear, and if one of them is an ordinary circle and the other an ordinary line, it is also clear, since no ordinary circle passes through the special point. If both conformal circles are lines, they have their ordinary point of intersection (if they are not parallel) and the special point as a point of intersection. If they are parallel, we say, very naturally, that they touch at the special point. Moreover, we can define the angles at which conformal circles meet. If the points of intersection are ordinary points, there is no difficulty; if the point of intersection is the special point ω, the two conformal circles are lines, and we define the angle at which they meet at ω to be the same as the angle at which they meet as ordinary lines and to be zero if they are parallel as ordinary lines. With all these explanations, we may now state:

1. Inversion is a conformal transformation of the conformal plane.
2. Inversion preserves conformal circles (that is, the image of a conformal circle under an inversion is another conformal circle).

We may read off the theorems of Section 12 immediately from this. For example, Corollary 2 to Theorem 2 says that the image under I of a line not through O is a circle through O. Now a line not through O is a conformal circle through ω but not through O, and by (1) its image will be a conformal circle through O but not through ω—that is, an ordinary circle through O. Theorem 1 reads that a conformal circle not through O or ω is mapped into a conformal circle not through O or ω, and this is obvious, since O and ω are mapped into each other. The reader should check all the other cases in the same way. We may regard the straight lines as circles of infinite radius with the special point as center, and in this case Corollary 2 to Theorem 1 still holds, if we interpret $1/0 = \infty$.

Definition. Let us say that α is a *circular transformation* of the conformal plane K if it is a one-one mapping of K onto itself under which the image of a conformal circle is a conformal circle.

We have shown that inversion is a circular transformation. Also every similarity transformation of E induces a circular transformation β if we agree that $\beta(\omega) = \omega$. For similarity transformations are one-one mappings of E onto itself that preserve circles and straight lines, and the result follows easily. It is clear that the set of all circular transformations of K is a group.

As an exercise (which we will use later), the reader may check that if I is the inversion with center O and coefficient r, and J is the inversion with center O and coefficient s, then IJ is the homothetic transformation with center O and coefficient s^2/r^2. What is the product of I with the homothetic transformation with center O and (positive) coefficient k?

We come now to the fundamental theorem on circular transformations:

Theorem 1. *Any circular transformation L of K either is induced by a similarity mapping, or can be represented as the product of a similarity and an inversion.*

Proof. If L leaves the special point ω fixed, then, confined to E, it maps lines into lines. For let l be any ordinary line and l' the conformal line corresponding to it (that is, l together with ω). Then L maps l' onto a conformal circle, and since $L(\omega) = \omega$, this conformal circle passes through ω, and so it is a line. The image of l under L confined to E is thus a line. Hence L is an affine transformation of E (when regarded as acting on E). For we have shown it preserves lines, and it is clearly one-one and onto. Now L maps ordinary circles into ordinary circles. Since it is an affine mapping, it can be represented as the product of an orthogonal transformation and two compressions in perpendicular directions (theorem in Section 26, Vol. 1). An orthogonal transformation takes a circle into a circle, and two perpendicular compressions take a circle into an ellipse whose major and minor axes are along the directions of the compression and the ratio of whose axes is the same as the ratio of the coefficients of compression. It follows that in this representation of L the coefficients are equal and hence that L is a similarity transformation, whose coefficient is precisely the common value of the coefficients of these two compressions (see the end of Section 26).

Thus a circular transformation leaving ω invariant induces a similarity transformation of E, and, conversely (as we noted before), every similarity transformation of E induces a unique circular transformation of K.

Suppose now that L takes ω into the point $O \neq \omega$. Let I be any inversion with pole O. It is certainly a circular transformation. So the product LI is a circular transformation, and since LI leaves ω fixed, it is an affine transformation A; that is,

$$LI = A;$$

$$L = L(II) = (LI)I = AI;$$

(since II is the identity transformation of K). We have thus represented L as the product of an inversion I and an affine transformation A. A and I need not commute. ▼

Note. The inversion I, like L, takes ω into O, so that A must leave O fixed. This means that A must be the product of a homothetic transformation with center O and either a rotation about O or a reflection in a line through O. (See Theorem 2, Section 17, in Volume 1.)

Theorem 2. *If the circular transformation L is not a similarity transformation, it can be represented in the form $L = \Omega I$, where Ω is an orthogonal transformation having a fixed point O, say, and I is an inversion with center O. A representation in such a form is unique, and, moreover, Ω and I commute—$\Omega I = I\Omega$.*

Proof. Since L is not a similarity transformation, it takes ω into some ordinary point O. We take our inversion I to have center O (as we obviously must).

Let M be any ordinary point whose image under L is not ω, and let M' be its image. We take the coefficient of I to be r, where $r^2 = OM \cdot OM'$. We have thus completely determined I, and we have

$$L = \Omega I,$$

where $\Omega = LI$ may be shown, just as in the proof of Theorem 1, to be a similarity transformation. Now if $I(M) = P$, then P lies on OM and

$$OP \cdot OM = r^2 = OM \cdot OM',$$

so that $OP = OM'$. But $\Omega(P) = M'$, $\Omega(O) = (O)$, so that the coefficient of Ω is 1, and Ω is orthogonal. In particular, since it fixes O, Ω must be a rotation about O or a reflection in a line through O or the identity.

To prove uniqueness, suppose also that $L = \Omega' I'$, where I' is an inversion with center O', say, and Ω' is an orthogonal transformation leaving O' fixed. Then $L = \Omega' I'$ takes ω into O', so that $O' = O$.

It follows from

$$\Omega I = \Omega' I'$$

that

$$\Omega'^{-1}\Omega = I'I.$$

Here the left side is an orthogonal transformation leaving O fixed, and the right side is the product of two inversions with coefficients r' and r, say, and centers O. By a result already left to the reader, this is a homothetic transformation with coefficient r'^2/r^2. Since it is also an orthogonal transformation, we must have $r' = r$, so that $I' = I$ (same coefficient and same center). Thus the right side above is the identity, and therefore the left is also, which means that

$$\Omega' = \Omega, \qquad I' = I.$$

It remains to show that $\Omega I = I\Omega$. This follows very easily from the fact that Ω is a rotation about O or a reflection in a line through it. We leave the details to the reader. ▼

Note. When we say in these theorems that a circular transformation of the conformal plane is the product of an inversion and a similarity mapping, or of an inversion and an orthogonal mapping, we mean, of course, a similarity (orthogonal) mapping as defined for the *conformal* plane. These are defined in the obvious way as those mappings whose restrictions to the ordinary part of the conformal plane are similarity (orthogonal) mappings and that take ω into itself. Similarly, when we say above that Ω is a rotation or a reflection, we mean that it induces a rotation or a reflection in the ordinary part of the plane.

Theorem 3. *If L is a circular transformation of the conformal plane K leaving fixed three distinct points A, B, C, then L is either the identity or the inversion with respect to the (unique) conformal circle passing through A, B, C.*

The reader should first verify that given any three distinct points of K, there exists exactly one conformal circle through them. This conformal circle will be a straight line if and only if either A, B, C are collinear (in the ordinary sense) or one of them is ω.

Proof. We use Theorem 2 to distinguish cases:

Case 1. L is a similarity. Since at least two of A, B, C are ordinary points, the restriction of L to E is a similarity fixing two distinct points; if L preserves orientation, it must be the identity (Volume 1, Section 15). If it reverses orientation, ABC cannot be a triangle (since if it were, L would preserve its orientation). Thus ABC is a straight line, and L must be the reflection in it.

We may regard this as a special case of inversion—that in which the circle of inversion is a straight line. In this case we must regard the center of the "inversion" $I = L$ as ω, so that ω is fixed under I. If we wish to extend the definition of inversion in this way, we must specify $O \neq \omega$ in Theorem 2 if we are to preserve the uniqueness part of that theorem.

Case 2. $L = \Omega I$ is the product of an inversion with center O ($\neq \omega$) and an orthogonal transformation Ω leaving O fixed. Since $L(O) = \omega$ and $L(\omega) = O$, A is not O or ω. If $I(A) = A'$, then $\Omega(O) = O$ and $\Omega(A') = A$, so that $OA' = OA$. This is only possible if A lies on the circle of inversion S. Repeating the argument with B and C, we see that S must be the circle ABC. Since I leaves S pointwise invariant, so does Ω, which means that Ω is the identity, and $L = I$ is the inversion in ABC. ▼

Theorem 4. *There exist two and only two circular transformations taking three given distinct points A, B, C into three given distinct points A', B', C'.*

Proof. Let I_1 be any inversion with center A, and let I_2 be any inversion with center A'. By this we mean that if A, for example, is an ordinary point, I_1 is to be an inversion with respect to some circle with center A. If $A = \omega$, I_1 is to be the reflection in some straight line (circle with center ω).

Let $I_1(B) = B_1$, $I_1(C) = C_1$, and $I_2(B') = B_2$, $I_2(C') = C_2$.

Let S be a similarity transformation (defined on K) taking the segment B_1C_1 onto the segment B_2C_2. Then it may be verified that $L = I_2SI_1$ takes A, B, C onto A', B', C', respectively. It is clear that L is a circular transformation.

Suppose now that L and L' both take A, B, C into A', B', C', respectively. Then $L^{-1}L'$ leaves A, B, C fixed, and so, by Theorem 3, it is the identity or the inversion I. Thus L' is either L or LI, where I is the inversion in the conformal circle ABC. ▼

APPENDIX 2 TO CHAPTER I

Principle of Duality

We recall from Section 9.2 (p. 83) the definition of dual statements in projective spaces. Closely connected with this concept is a variety of general propositions which together or separately are referred to as the *principle of duality*.

Suppose we are given an abstract projective space Σ, and a statement A about Σ. We say A is a *theorem* of Σ if there exists a *proof* of A. A proof P of A is in turn defined as follows: P is a sequence of statements of which the last is A, such that each statement either is an axiom for Σ (a statement which is "given" as true in Σ) or follows immediately from previous statements of P by the laws of logic. It is clear that this conforms to our intuitive idea of what a proof should be.

Suppose now for simplicity that the only axioms that occur in P are those that we listed above for any abstract projective space (axioms I–VIII, and L). Then P will contain only *primitive* notions: those of point, line, plane, and incidence. We may therefore write the dual sequence of statements P^*. It is clear that each statement of P^* either is the dual of an axiom or follows logically from previous statements of P^*. Now the duals of all the axioms are theorems: those of axioms I–VIII are again axioms, and we have already proved the dual L' of L. It follows at once that P^* can be expanded to a proof of its final statement, that is, of A^*, by inserting a proof of L' at the first occasion where P refers to L.

Since P used only axioms that are true in *any* projective space (by definition of an abstract projective space), P^* will be a proof of A^* in *any* projective space, not merely in Σ. We thus have our first version of the principle of duality:

Theorem I. *If the statement A is a theorem for every abstract projective space, then so is A^*.*

As an example of such a statement A we give the following: If one line of Σ contains exactly $q + 1$ points (for some integer q), then every line contains $q + 1$ points, every plane contains $q^2 + q + 1$ points, and Σ itself contains $q^3 + q^2 + q + 1$ points. The reader is invited to write the dual statement A^*, and to prove A. A projective space having the property that some line has a finite number of points is called a *finite* projective space. It can be proved that for such a space q must be a prime power: $q = p^r$. Conversely, there is exactly one abstract space having a given number $q + 1 = p^r + 1$ of points on each line. Clearly, no such space is isomorphic to real projective space.

So far we have considered only statements A that hold in *all* projective spaces; so long as we deal with only such statements we are still at the foundations of the subject. In particular projective spaces, for example the real projective space, there may be a great number of important theorems which will not be true, or even meaningful, in *every* projective space. For example, it is possible to define an ordering of the points of every line in real projective space in such a way that a number of pleasant properties hold, but there are many projective spaces in which this is impossible. A proof P of a statement A in a particular projective space Σ (say the real one) may well contain axioms for Σ which are not true (or, perhaps, meaningful) in every space. In general, before we can talk of a dual theorem A^* there are thus two difficulties to overcome.

First, we may not know how to write the dual of A at all, because A may contain notions other than the primitive ones, and we may not know how to define *their* duals. Let us avoid this difficulty for a moment, by supposing that both A and a

proof P for A contain only primitive notions. Then we may write the dual sequence of statements P^*, and the last statement of P^* will be A^*.

We now come to the second difficulty: there is no reason why P^* should be a proof of A^*. P^* will be a sequence of statements such that each either follows logically from previous statements of P^*, or is the dual of an axiom for Σ. The reason why P^* need not lead to a proof of A^* is that the dual of an axiom for Σ need not be a theorem of Σ. To deal with this situation, we introduce a new concept.

Suppose Σ is any abstract projective space. We define the *dual space* Σ^* of Σ as follows. The points of Σ^* will be objects in one-one correspondence with the planes of Σ. Let us refer to them as "points." Then we define three "points" to be collinear if and only if the corresponding planes of Σ are collinear (all incident to a single line). Thus a line of Σ^* (a "line") will consist of all those "points" such that the corresponding planes of Σ are collinear. Finally, the planes of Σ^* ("planes") will consist of all those "points" such that the corresponding planes of Σ are concurrent in some point of Σ. Thus we have a natural way of identifying the lines of Σ^* with the lines of Σ and the planes of Σ^* with the points of Σ, as well as the points of Σ^* with the planes of Σ. Note that a line of Σ, when considered as a line of Σ^*, must be thought of as consisting of the set of all the planes through it (like the spine of a book), rather than all the points on it. Similarly, a point of Σ, when identified with a plane of Σ^*, must be thought of as the collection of all the planes of Σ through it (a "star" of planes).

To check that Σ^* with these definitions of "point," "line," and "plane" is a projective space, we must verify the axioms. It will be found that each axiom for Σ^* is just a restatement of the dual of that axiom in Σ. Since the dual of an axiom of Σ is either an axiom (for the axioms I–VIII) or a theorem (for L), it follows that the axioms for a projective space do indeed hold in Σ^*.

As an exercise the reader may like to consider what the dual space of Σ^* is; it turns out that it may be naturally identified with Σ: $\Sigma^{**} = \Sigma$ in the sense of isomorphism.

Now consider a statement A about a space Σ in which only the primitive notions occur. We may think of the process of dualizing A as a process of *coding*, provided that we understand the dualized statement to refer to Σ^*. For example, if A is the statement, "The line m (of Σ) is incident to exactly three points," the coding would go "The 'line' 'm' (of Σ^*) is incident to exactly three 'planes.'" This is just a disguised or coded way of saying exactly the same thing, as soon as we have identified the points, lines, and planes of Σ^* with the planes, lines, and points of Σ. Thus if A is true in Σ, then *by the very way we have defined* Σ^*, the coded statement (which is, of course, simply A^*) is a true statement *in* Σ^*.

We now return to the situation where A is a theorem for Σ having a proof P in which only primitive notions occur. In particular, the only axioms for Σ which occur in P are *primitive axioms*: axioms in which only primitive terms occur. Let us restrict our attention slightly further, and suppose that *every* axiom for Σ is primitive. Then we have:

Theorem 2. *Let Σ be a projective space having only primitive axioms. If A is a theorem of Σ, then A^* is a theorem of Σ^*.*

This theorem, like Theorem 1, is called the principle of duality. The reader should check that Theorem 2 is stronger than Theorem 1. To do this, he will need to know that every space is (isomorphic to) the dual of some space. The exercise above settles this point, for it shows that any space is the dual of its own dual.

Proof of Theorem 2. If A is a theorem of Σ, there exists a proof P of A. It is clear from our hypothesis and the definition of a proof that P is a sequence of primitive statements. Thus we may write the dual sequence P^*. Each statement of P^* either follows logically from previous statements of P^*, or is the dual of an axiom for Σ. Each of the latter is a "disguised" axiom for Σ, thus true in Σ^*. It follows that P^* can be expanded to a proof in Σ^*; in fact, a proof of its last statement, which is just A^*. So A^* is a theorem of Σ^*. ▼

It may be noted that it is not necessary that the only *words* (apart, of course, from the logical and grammatical framework) appearing in the axioms of Σ, or in A, or in a proof P for A, be primitive. It is enough if the notions can be *translated back* to primitive notions. For example, the word "quadrangle" is "effectively" primitive, as is the relation "is in perspective with" (of two quadruples of collinear points).

To say that Σ has only primitive axioms is to say that Σ is a member of the class of those projective spaces which satisfy a certain list of primitive properties (the axioms). As a well-known example of such a class of spaces defined by primitive properties we may cite the *Desarguesian* spaces: those spaces which satisfy, in additional to the axioms I–VIII and L, an axiom stating that the *Desargues property* holds: Whenever ABC and $A'B'C'$ are triangles such that AA', BB', CC' are concurrent, it follows that aa', bb', cc' are collinear, where aa' is the point of intersection of $BC = a$ and $B'C' = a'$, and so on.

It is not very difficult to show that the class of Desarguesian spaces coincides with the class of *all* spaces: that is, the Desargues property is a theorem for any projective space.

We now come to the first difficulty we raised over dualizing a theorem A of Σ: A may contain concepts for which we do not know the duals. There is nothing to be said about this in general; we must consider in each case how to define the dual of a concept so that statements in which that concept occurs have "satisfactory" dual statements. A "satisfactory" dual statement is, of course, one such that Theorem 2 will continue to hold good. There is no guarantee, in general, that a given concept *has* a satisfactory dual.

We confine out attention here to one important special case: real projective space and the important, but nonprimitive, notion of cross ratio. It turns out that the "correct" way to define the dual of "cross ratio" is "cross ratio." Note that this definition makes sense: that is, the dual of a sentence in which the word "cross ratio" appears will be meaningful. This is because we have taken care to define cross ratio not only on

quadruples of collinear points, but also on the dual configuration: quadruples of collinear planes. We now state without proof our justification for this definition.

If A is a theorem of real projective space Σ containing only primitive notions and the notion of cross ratio, then the dual statement A^* is a theorem of Σ^*.

Once again, A can contain further notions, provided that they can be reduced to these. For example A is allowed to contain the words "a harmonic set."

We saw in Section 9.3 that the dual Σ^* of real projective space Σ (and similarly the dual of complex projective space) is isomorphic to Σ itself. Note that this is a special result for these special spaces; it must not be confused with the general result, for any projective space Σ, that Σ and Σ^{**} are isomorphic. Thus we have the following principle of duality for real (and complex) projective space:

Theorem 3. *If* A *is a theorem of real (or complex) projective space* Σ, *and* A *contains only primitive notions and the notion of cross ratio, then* A^* *is also a theorem of* Σ.

So far we have been talking about only projective *spaces*. There is also a variety of duality principles for projective *planes*, in which we interchange "line" and "point" to obtain the dual of a primitive statement. Almost all our discussion has obvious (and true) analogs for this situation: the analog of Theorem 1, the definition of the *dual plane* Π^* of a given abstract projective plane Π, the analog of Theorem 2, the dual of "cross ratio," and the analog of Theorem 3.

As examples, the reader is invited to write the dual statement of the Desargues property (which still makes sense in the plane), to verify that the dual of a quadrangle is a quadrilateral, and to check that in the real projective plane the theorems of Section 5.3 are dual to each other.

The analog of the statement A given as an example to Theorem 1 is true; the reader may write and prove it. However,

it is not known whether q must be a prime power (this is a famous open question), and whereas for every $q = p^r$ there exists at least one projective plane with $q + 1$ points on each line, there exist values of q for which there is more than one such plane (in the sense of isomorphism).

As we mentioned, the statement of the Desargues property still makes sense in a projective plane, and we may speak of a *Desarguesian plane*, meaning a plane in which the property always holds. We could also define the class of Desarguesian planes to be the class of those projective planes which satisfy, in addition to the three axioms for any projective plane (Section 1.6) the (primitive!) axiom stating that the Desargues property holds.

It is not true that every projective plane is Desarguesian; in fact we have the following result: *The fundamental theorem* (Section 3.1) *holds in a projective plane* Π *if and only if* Π *is Desarguesian.*

Subject Index

tr. = transformation(s).

References in parentheses are to passages in which either the actual *words* of the entry do not occur, or the words have a meaning other than their main meaning.